

# **PROYECTO FIN DE MÁSTER EN INGENIERÍA INFORMÁTICA PARA LA INDUSTRIA**

**MÁSTER EN INVESTIGACIÓN  
EN INFORMÁTICA  
Facultad de Informática  
Universidad Complutense de Madrid**

## **ANÁLISIS DE MÉTODOS DE PROCESAMIENTO DE IMÁGENES ESTEREOSCÓPICAS FORESTALES**



Alumno:  
Álvaro Suárez Bravo

Director:  
Gonzalo Pajares Martinsanz

**2008-09**



# Resumen

Mediante imágenes de zonas boscosas se puede determinar el volumen de madera presente en la imagen. Para ello se cuenta con un sistema de visión estero. El proceso puede dividirse en dos partes, la segmentación de las regiones correspondientes a los troncos y la reconstrucción del escenario 3D para el cálculo del volumen de madera proporcionado por los mismos.

Para la segmentación se estudiarán técnicas de segmentación por umbralización y segmentación por detección de regiones. Para la reconstrucción del escenario se estudiarán distintos algoritmos para resolver el problema de la correspondencia. La parte de triangulación basada en fundamentos geométricos, necesaria para el cálculo de las coordenadas de un punto en el espacio a partir de la proyección de ese punto en las dos imágenes, no se aborda en este documento.

**Palabras clave:** imágenes forestales, segmentación, texturas, umbralización, regiones, visión estéreo, correspondencia, triangulación.

# Abstract

By means of images of wooded zones it is possible to determine the volume of wood existing in the image by using a stereo vision system. The process can be divided in two parts, the segmentation of the regions corresponding to the trunks and reconstruction of the 3D scene to calculate the volume of wood supplied by them.

For segmentation stage, threshold segmentation and region detection techniques will be studied. For reconstruction stage different algorithms will be studied to solve the correspondence problem. Triangulation based on geometric fundamentals which is

necessary to calculate the coordinates of a point in space from the projection of that point in the two images, is not discussed in this document.

**Keywords:** forest images, segmentation, textures, thresholding, regions, stereo vision, matching, triangulation.

# Agradecimientos

A Fernando Pita Ingeniero de Montes y a la Dra. Isabel Cañellas quienes a través del Centro de Investigación Forestal (CIFOR) del Instituto Nacional de técnicas Agroalimentaria (INIA) han facilitado las imágenes estereoscópicas con las que se ha realizado este trabajo.



# Autorización

El abajo firmante, matriculado/a en el Máster en Investigación en Informática de la Facultad de Informática, autoriza a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a su autor el presente Trabajo Fin de Máster: “Análisis de métodos de procesamiento de imágenes estereoscópicas forestales”, realizado durante el curso académico 2008-2009 bajo la dirección del Profesor Gonzalo Pajares Martinsanz en el Departamento de Arquitectura de Computadores y Automática, y a la Biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en Internet y garantizar su preservación y acceso a largo plazo.





# Índice

<b>1. INTRODUCCIÓN.....</b>	<b>11</b>
1.1. IDENTIFICACIÓN DEL PROBLEMA .....	11
1.2. MOTIVACIÓN Y OBJETIVOS DE INVESTIGACIÓN .....	13
1.2.1. Motivación.....	13
1.2.2. Objetivos generales .....	14
1.2.3. Objetivos específicos.....	14
1.3. APORTACIONES A LA INVESTIGACIÓN .....	14
<b>2. REVISIÓN DE MÉTODOS .....</b>	<b>17</b>
2.1. INTRODUCCIÓN.....	17
2.2. SEGMENTACIÓN.....	17
2.2.1. Segmentación basada en la umbralización por histograma .....	18
2.2.2. Segmentación basada en regiones .....	20
2.2.2.1. Algoritmo de crecimiento de regiones .....	22
2.2.2.2. Algoritmo de División y fusión (“Split & Merge”).....	25
2.2.2.3. Watershed.....	27
2.3. VISIÓN ESTÉREO .....	28
2.3.1. Problema de la correspondencia.....	29
2.3.1.1. Métodos basados en el área .....	30
2.3.1.2. Métodos basados en las características.....	32
2.3.1.3. Métodos basados en el área frente a métodos basados en características.....	35
<b>3. CLASIFICACIÓN DE TEXTURAS Y CORRESPONDENCIA ESTEREOSCÓPICA.....</b>	<b>37</b>
3.1. ANÁLISIS DE LAS IMÁGENES: PROBLEMÁTICA.....	37
3.2. SEGMENTACIÓN.....	40
3.2.1. Segmentación basada en la umbralización por histograma .....	40
3.2.2. Segmentación basada en regiones .....	43
3.2.2.1. Algoritmo de crecimiento de regiones .....	43
3.2.2.2. Algoritmo de División y Fusión “Split&Merge”.....	47
3.2.2.3. Watershed.....	52
3.2.2.4. Resumen de los métodos de segmentación.....	53

3.3.	CORRESPONDENCIA .....	53
3.3.1.	Métodos basados en el área .....	58
3.3.2.	Métodos basados en características .....	65
3.3.3.	Resumen de los métodos de correspondencia .....	67
<b>4.</b>	<b>ANÁLISIS DE RESULTADOS .....</b>	<b>69</b>
4.1.	OBJETIVOS DEL ANÁLISIS.....	69
4.2.	ANÁLISIS DE RESULTADOS.....	71
4.2.1.	Criterio de validación .....	71
<b>5.</b>	<b>CONCLUSIONES Y TRABAJO FUTURO.....</b>	<b>73</b>
5.1.	SOBRE RESULTADOS .....	73
5.2.	CUMPLIMIENTO DE LOS OBJETIVOS PROPUESTOS .....	74
5.3.	TRABAJO FUTURO .....	74
<b>6.</b>	<b>BIBLIOGRAFÍA.....</b>	<b>75</b>

# Capítulo 1

## 1. Introducción

### 1.1. Identificación del problema

Es preciso reconocer que hoy por hoy la visión por computador a veces no es la mejor solución a un problema. Existen muchas ocasiones en las que el problema es tan complejo que la solución humana sigue siendo lo ideal. Por ejemplo imaginemos una conducción de un vehículo en una carretera con tráfico intenso. A veces, las soluciones humanas tienden a ser inexactas o subjetivas y en ocasiones lentas y presentan una ausencia de rigor así como una pobre percepción (Marshall y Martin 1993). No obstante, la solución humana es menos estructurada que la solución artificial y muchos problemas de visión por computador requieren un nivel de inteligencia mucho mayor que el que la máquina pueda ofrecer. El sistema de visión humana puede describir automáticamente una textura en detalle, un borde, un color, una representación bidimensional de una tridimensional, ya que puede diferenciar entre imágenes de diferentes personas, firmas, colores, etc., puede vigilar ciertas zonas, diagnosticar enfermedades a partir de radiografías, etc. Sin embargo, aunque algunas de estas tareas pueden llevarse a cabo mediante visión artificial, el software o el hardware necesario no consigue los resultados que serían deseables.

Aún a pesar de las limitaciones expuestas, cada día es mayor el número de aplicaciones de la visión artificial. Por ello sólo es posible dar una pequeña pincelada sobre las múltiples aplicaciones en las que la visión artificial se ha aplicado hasta el momento. Aquí solamente proporcionamos algunas ideas para generar aplicaciones en la mente del lector. Obviamente los ejemplos que se presentan son aplicables a cualquier otro proceso diferente y en otro campo industrial o científico también diferentes. En Newman y Jain (1995) o Davies (1997) se puede encontrar una amplia revisión sobre la inspección visual y aplicaciones de la visión a diferentes campos de la vida real, que nos proporcionan un amplio abanico de posibilidades. Para dar idea tan sólo algunas de tales aplicaciones:

- 1) Control de cultivos en agricultura, propiciado por la necesidad de conocer los cultivos programados para la recepción de subvenciones u otros aspectos relacionados.
- 2) Cómputo y medición de parcelas agrícolas y tipo de cultivos a los que se dedica.
- 3) Control de riegos agrícolas
- 4) Agricultura de precisión para aplicar herbicida en el tratamiento de malas hierbas de forma selectiva, evitando la contaminación medioambiental y la reducción de costes de producción.
- 5) Evaluación de catástrofes naturales: fuegos, daños por inundaciones, heladas en cultivos agrícolas, nevadas, etc.
- 6) Detección de cambios en determinadas zonas, principalmente urbanas para el control de edificaciones o impactos medioambientales
- 7) Vigilancia en prevención de catástrofes, por ejemplo fuegos o inundaciones
- 8) Control de fenómenos meteorológicos como es el retroceso de determinadas playas
- 9) Vigilancia: forestal, marítima.
- 10) Detección de infraestructuras: carreteras, caminos forestales, cañadas reales, etc.
- 11) Finalmente, destacamos la aplicación objeto de esta memoria, que en sí misma constituye una aplicación novedosa al tratarse de la obtención del volumen de madera en una zona boscosa.

Sobre esta última aplicación y hasta donde se ha podido investigar, no existe una base bibliográfica de trabajos previos específicos que permitan establecer un punto de partida concreto y preciso.

El trabajo de investigación se aborda desde dos perspectivas, por un lado la clasificación de texturas y por otro la correspondencia estereoscópica propiamente dicha. La primera tiene como finalidad extraer aquellas partes de la imagen de interés como son los troncos e identificar aquellas otras que carecen del mencionado interés por no ser portadoras de madera, como es el caso de las texturas correspondientes a las hojas, suelo, cielo, etc. Identificadas las texturas, el siguiente paso consiste en tratar de emparejar un tronco de una

imagen con su homólogo en la otra imagen del par estereoscópico. Estos son los dos aspectos abordados desde diferentes perspectivas.

La presente memoria se organiza en cinco capítulos. En lo que resta del presente capítulo se exponen los objetivos que se plantearon en la investigación aquí recogida, en base a los planteamientos previos y que tratan de dar solución a la problemática existente. También se exponen las estrategias realizadas para abordarlos. En el capítulo 2 se realiza una revisión del estado del arte en el tema de clasificación de texturas así como en lo que respecta a métodos de correspondencia. Ambos aspectos constituyen la base del trabajo de investigación. En el capítulo 3 se propone la estrategia de solución a la problemática planteada, junto con los pasos del proceso que la conforman. En el capítulo 4 se realiza un análisis sobre los resultados obtenidos mediante las estrategias planteadas. Finalmente en el capítulo 5 se extraen las conclusiones pertinentes y se exponen las líneas de investigación futuras, que dan pie al inicio de la investigación en los estudios de Doctorado.

## **1.2. Motivación y objetivos de investigación**

### **1.2.1. Motivación**

Como se ha mencionado anteriormente, la investigación consiste en la identificación de las texturas existentes en zonas boscosas con el fin último de estimar el volumen de madera existente en una determinada zona a partir de un sistema de Visión estereoscópica, que requiere a su vez el emparejamiento de los árboles en las dos imágenes del par estereoscópico.

Este trabajo tiene su origen en el proyecto de investigación titulado “Análisis estereoscópico y reconocimiento de texturas en imágenes forestales obtenidas mediante dispositivo de medición MU-200501738”. Dicho proyecto se desarrolla en colaboración con el Instituto Nacional de Investigación Agraria y Agroalimentaria (INIA) y la Escuela Técnica Superior de Ingenieros de Montes (ETSIM) de la Universidad Politécnica de Madrid siendo investigador principal del mismo el director de este trabajo de investigación.

La captura de las imágenes se lleva a cabo en el bosque objeto de análisis, situándose en diferentes áreas hasta completar la superficie total bajo análisis. Actualmente, las imágenes se procesan manualmente en los laboratorios de la ETSIM o del INIA para estimar el

volumen de madera, siendo el objeto de este trabajo, la investigación de técnicas con el objetivo de automatizar el proceso de estima.

### **1.2.2. Objetivos generales**

- 1) Aprender a manejar referencias bibliográficas, así como la forma de abordar las investigaciones
- 2) Aprender y familiarizarse con la mayoría de técnicas existentes de visión por computador.
- 3) Analizar el problema y ver cuáles de los métodos anteriores o combinaciones de los mismos podrían aplicarse a la resolución del problema de la estima automática de la madera existente en el área identificada por el par de imágenes estereoscópicas.
- 4) Analizar las ventajas e inconvenientes de cada uno de ellos así como su problemática
- 5) Dar solución a un problema realizando un aporte de carácter investigador.

### **1.2.3. Objetivos específicos**

- 1) Identificar y diseñar un algoritmo mediante el cual se pueda dar una solución lo más precisa posible al problema.
- 2) Identificar una posible mejora de los métodos existentes o dar una solución novedosa a la problemática de la clasificación de las texturas naturales.
- 3) Implementar la mejora propuesta con el fin de analizar su comportamiento en base a los resultados obtenidos.
- 4) Identificar líneas de investigación futuras
- 5) Difundir los resultados mediante publicaciones científicas u otro tipo de divulgación

## **1.3. Aportaciones a la investigación**

Como resumen de todo lo anterior las aportaciones de investigación realizadas en el presente trabajo son las siguientes:

- 1) Diseño de métodos de detección de los troncos de los árboles
- 2) Diseño e implementación de algoritmos para resolver el problema de la correspondencia.
- 3) Algoritmos para la reconstrucción de la escena.
- 4) Estudiar la precisión y bondad de los resultados obtenidos.
- 5) Sentar las bases de trabajos futuros, y proponer futuras vías de investigación en base a los resultados obtenidos.





## 2. Revisión de métodos

### 2.1. Introducción

Siguiendo los pasos de investigación descritos en la sección 1.3 en este capítulo se describen los métodos investigados en relación a los dos objetivos planteados:

- Segmentación
  - Segmentación por umbralización
  - Segmentación por detección de regiones
- Visión Estereoscópica
  - Algoritmos para resolver el problema de la correspondencia
  - Reconstrucción de la escena

### 2.2. Segmentación

Básicamente, en cualquier imagen se encontrarán presentes uno o varios objetos localizados en un entorno. El objetivo de la segmentación consiste en separar dichos objetos del medio en el que se encuentran y distinguirlos entre sí. En el caso de las imágenes objeto de la investigación este esquema no resulta tan simple, antes bien presenta una enorme cantidad de texturas que dificultan su extracción, en cualquier caso, la segmentación, tanto en su versión simple como compleja se basa en los siguientes principios:

- Similitud: cada uno de los píxeles de un elemento tiene valores parecidos con respecto a alguna propiedad determinada.
- Discontinuidad: los objetos y más concretamente las regiones que los diferencian destacan del entorno y tienen por tanto unos bordes definidos.

- **Conectividad:** los píxeles pertenecientes al mismo objeto o región tienen que ser contiguos, es decir, tienden a agruparse constituyendo regiones homogéneas.

En nuestro caso, los elementos que queremos separar son los troncos de los árboles. Las imágenes que tenemos presentan ciertas dificultades debidas, en gran medida, a la iluminación. Los árboles no siempre poseen un color o textura que los diferencie fácilmente del resto de elementos. Debido también a la iluminación y al resto de elementos en la imagen, pueden aparecer regiones aparentemente no conectadas.

Tras el proceso de segmentación se espera tener una serie de regiones que definan los troncos cercanos al punto de vista. A continuación se describen una serie de métodos clásicos, que constituyen la base para realizar la segmentación.

### **2.2.1. Segmentación basada en la umbralización por histograma**

El histograma de una imagen nos permite conocer la frecuencia relativa de aparición de cada uno de los posibles niveles de intensidad dentro de la imagen en cuestión. El histograma nos servirá para agrupar los píxeles en función de su nivel de intensidad luminosa. En la figura 2.1 podemos ver una imagen muy sencilla formada por un objeto oscuro sobre un fondo gris uniforme. Su histograma se muestra en la figura 2.2.

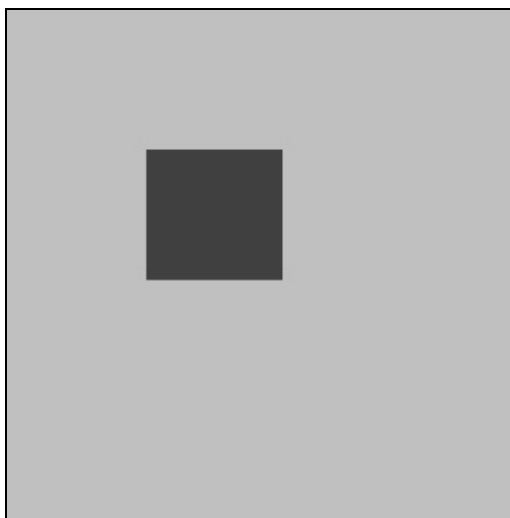


Figura 2.1 Imagen de color gris oscuro sobre un fondo gris claro.

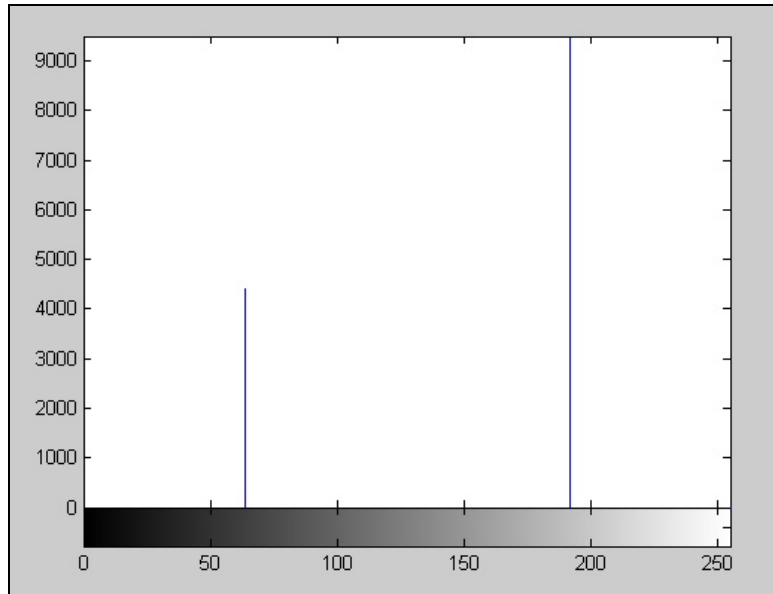


Figura 2.2 Histograma de la imagen de la Figura 2.1.

Puede verse en el histograma cómo aparecen dos líneas bien diferenciadas, que identifican las regiones existentes en la imagen, en este caso concreto y dada la imagen tan sencilla, se trata de un histograma bimodal. En este caso resulta fácil calcular un umbral para distinguir las dos clases.

La umbralización basada exclusivamente en los niveles de luminosidad se denomina global, mientras que en el caso de la umbralización local se utiliza información puntual de cada píxel, es decir, propiedades de un píxel que dependen de su localización en la imagen.

Existen dos métodos básicos para abordar el problema de la umbralización global (Maravall, 1993):

- Basada en la búsqueda de mínimos.
- Basada en las técnicas de reconocimiento de patrones.

La umbralización basada en la búsqueda de mínimos trata de encontrar niveles de intensidad que umbralicen el histograma de una imagen. Por otro lado, la umbralización basada en las técnicas de reconocimiento de patrones consiste, precisamente, en buscar patrones o formas en el histograma de la imagen.

En nuestro caso, nos centraremos en el primer tipo de umbralización.

### 2.2.2. Segmentación basada en regiones

La segmentación basada en regiones busca dividir la imagen en particiones que comparten ciertas características comunes. En la figura 2.3 podemos ver una imagen en la que se han identificado tres regiones diferenciadas por su color. La región 1 pertenecería al fondo mientras que las regiones 2 y 3 pertenecerían a elementos de la imagen.

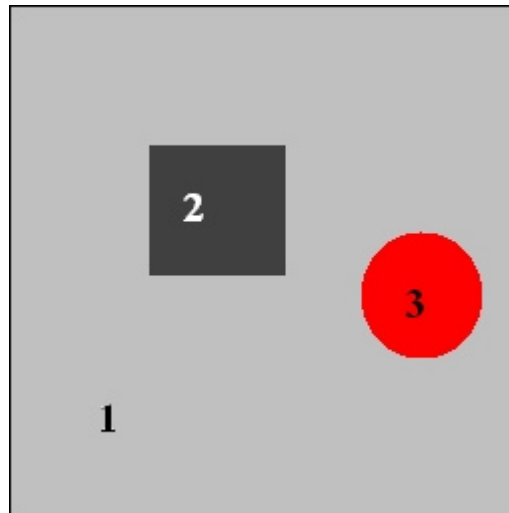


Figura 2.3 Imagen en la que aparecen dos objetos (2 y 3) sobre fondo gris claro.

El algoritmo general en la segmentación por detección de regiones consta de 3 etapas:

1. Iniciar la partición
2. Definir un criterio de homogeneidad
3. Optimizar la partición mediante modificaciones (división, unión, etc.). En algunos casos después de esta etapa volveremos de forma iterativa al paso 2.

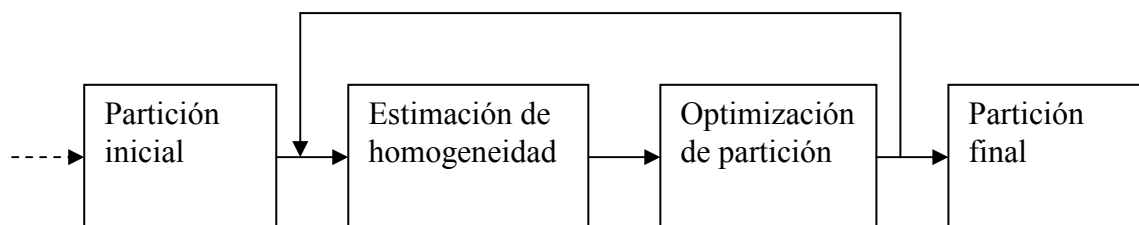


Figura 2.4. Etapas del algoritmo de segmentación basada en regiones.

1. Iniciar la partición. En esta primera etapa realizamos una primera segmentación muy simple. Una posible inicialización puede ser la existencia de una única región

que engloba a toda la imagen, o justamente lo contrario, una región por cada píxel que forma la imagen. Esta inicialización también dependerá del conocimiento que tengamos sobre las regiones a extraer y de otros factores como puede ser la supervisión por parte de un experto que elija puntos clave para la posterior segmentación.

2. Criterio de homogeneidad. A la hora de segmentar una imagen debemos seleccionar unos criterios que nos permitan clasificar unas regiones como pertenecientes a una clase u otra. El color y las texturas son dos de estas características. Puede considerarse que un píxel forma parte de una región cuando comparte propiedades comunes con el resto de píxeles de dicha región. En ocasiones la pertenencia o no a una región no debe tratarse a un nivel tan bajo como un píxel sino a un conjunto de ellos, como ocurre con las texturas de las regiones o los contornos que las definen.

Existen dos alternativas para aplicar el criterio de homogeneidad: determinista y aleatoria (Trucco y Verri, 1998).

- Determinista:
  - Textura: Varianza respecto a un modelo.

$$C_T = \sum_n \sum_{i,j \in R_n} [x(i,j) - M_n(i,j)]^2 \quad (2.1)$$

Siendo R la región, (i, j) la posición y M el modelo con el que se quiere comparar.

- Contorno: minimizar la longitud para tener contornos simples. Un contorno simple es aquél que bordea la región objeto de estudio adaptándose a ella mediante la mínima longitud posible. Si podemos describir el contorno de una región podremos posteriormente comparar regiones a partir de estos descriptores de contorno.

$C_c$  = Longitud de contornos

$$C = aC_t + (1 - a)C_c \quad (2.2)$$

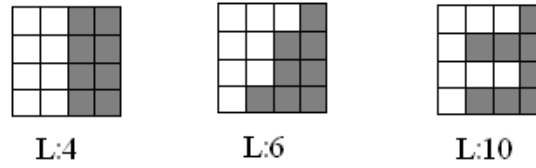


Figura 2.5 Ejemplo de cálculo de la longitud de Contornos.

- Aleatorio: estudio del contorno de configuraciones locales.

$$C = C_T \prod_{\text{Transiciones}} e^{(\nu_c)} \quad (2.3)$$

3. Modificaciones de la partición: en la última etapa del algoritmo, las particiones son modificadas y actualizadas para mejorar las condiciones de homogeneidad. Las principales modificaciones que se realizan en esta fase consisten en la división y unión de regiones.

Dentro de la segmentación basada en regiones vamos a estudiar tres algoritmos clásicos:

1. Algoritmo de crecimiento de regiones.
2. Algoritmo de Fusión y División “Split & Merge”.
3. Watershed.

Podemos encontrar el algoritmo de crecimiento de regiones y el de fusión y división en Fu y col. (1988), Sonka y col. (1995) o González y Woods (1993). En Coiras y col. (1998) se puede encontrar un método de crecimiento de regiones. El método de Watershed fue originalmente propuesto por Digabel y Lantuéjoul (1978) y posteriormente mejorado por Beucher and Lantuéjoul (1979).

#### 2.2.2.1. Algoritmo de crecimiento de regiones

La segmentación por crecimiento de regiones se basa en dos de los principios anteriormente mencionados: proximidad y similitud de puntos.

Para cada región que se desea segmentar, el crecimiento se inicia a partir de uno o más puntos iniciales, conocidos como semillas. A partir de este momento, la región va creciendo e incorporando a aquellos vecinos inmediatos que satisfacen un criterio de aceptación prefijado. El crecimiento de esta región termina cuando no hay elementos que

cumplan tal condición. Al finalizar el proceso tendremos un conjunto de regiones que engloban a un conjunto de puntos que la definen.

### ***Selección de semillas***

El principal problema en el caso del crecimiento de regiones es la ubicación de las semillas sobre la imagen de forma que a partir de esas semillas se generen las regiones. Por ello es importante que las semillas se especifiquen en posiciones de la imagen que corresponden a puntos de interés. La selección de las semillas puede hacerse de forma automática o con la ayuda de un experto que las seleccione de forma manual, que sería otra posible opción para que el método se orientara de una forma semiautomática.

### ***Criterio de aceptación de los puntos***

Se debe establecer un criterio apropiado para extraer las componentes de interés a partir del conjunto de puntos (píxeles) iniciales, definiendo las propiedades que dichos puntos deben cumplir para resultar incorporados a la región. Este criterio normalmente se basa en la proximidad y la homogeneidad de los píxeles adyacentes.

Por lo tanto, la elección del criterio a emplear para la evaluación de puntos candidatos constituye un aspecto crítico para el adecuado funcionamiento del algoritmo.

Un criterio simple de comparación se basa en el análisis del valor de intensidad del punto candidato con respecto al de las semillas especificadas, obteniendo una medida de distancia como sigue

$$d(v) = |I(v) - \bar{I}(s)| \quad (2.4)$$

donde  $I(v)$  es el valor de intensidad del punto candidato  $v$  e  $I(s)$  se refiere a la intensidad de cada uno de los puntos  $s$  del conjunto  $S$  de semillas especificadas para la región. Además pueden considerarse los valores  $d(v')$  correspondientes a los puntos  $v'$  en el entorno de  $v$ . Si para un determinado porcentaje  $p$  de los puntos evaluados se satisface que el valor  $d$  se encuentra dentro de un cierto límite de tolerancia  $t$ , el punto  $v$  se integra a la región, de lo contrario se descarta. Esta consideración permite lograr un criterio más robusto de crecimiento, el cual contribuye a prevenir situaciones de desbordamientos hacia otras componentes por conductos delgados y también problemas de sub-segmentación,

generalmente debido al ruido. El hecho de considerar los puntos  $v'$  también conlleva un procesamiento mas costoso.

Un aspecto importante del algoritmo es la determinación de un valor de tolerancia para la diferencia de intensidades de los puntos de la región. El concepto de homogeneidad es bastante impreciso para el usuario, generalmente depende de las características de la imagen y varía según las distintas modalidades, por lo que no siempre es traducible efectivamente a un criterio computable, su especificación puede transformarse en un proceso de prueba y error.

### ***Estrategia de crecimiento***

La incorporación de puntos se basa en una estrategia de búsqueda en amplitud (Best First Search, BFS). Los puntos que son visitados y cumplen con la condición de aceptación son insertados en una lista hasta el momento de ser procesados y eventualmente integrados a la región. Inicialmente la lista está formada por las semillas. Luego se realiza el crecimiento en forma iterativa, extrayendo en cada ciclo el primer elemento  $u$  de la lista, el cual se etiqueta como perteneciente a la región para evitar su re-evaluación, y a partir de él se evalúan sus vecinos  $v$  en el entorno inmediato. Este esquema asegura que cada punto se evalúa en una única oportunidad, aunque sí es posible que sea considerado más de una vez en el análisis del entorno de sus vecinos. El pseudocódigo de la figura 2.6 resume el esquema del algoritmo de segmentación por crecimiento de regiones.



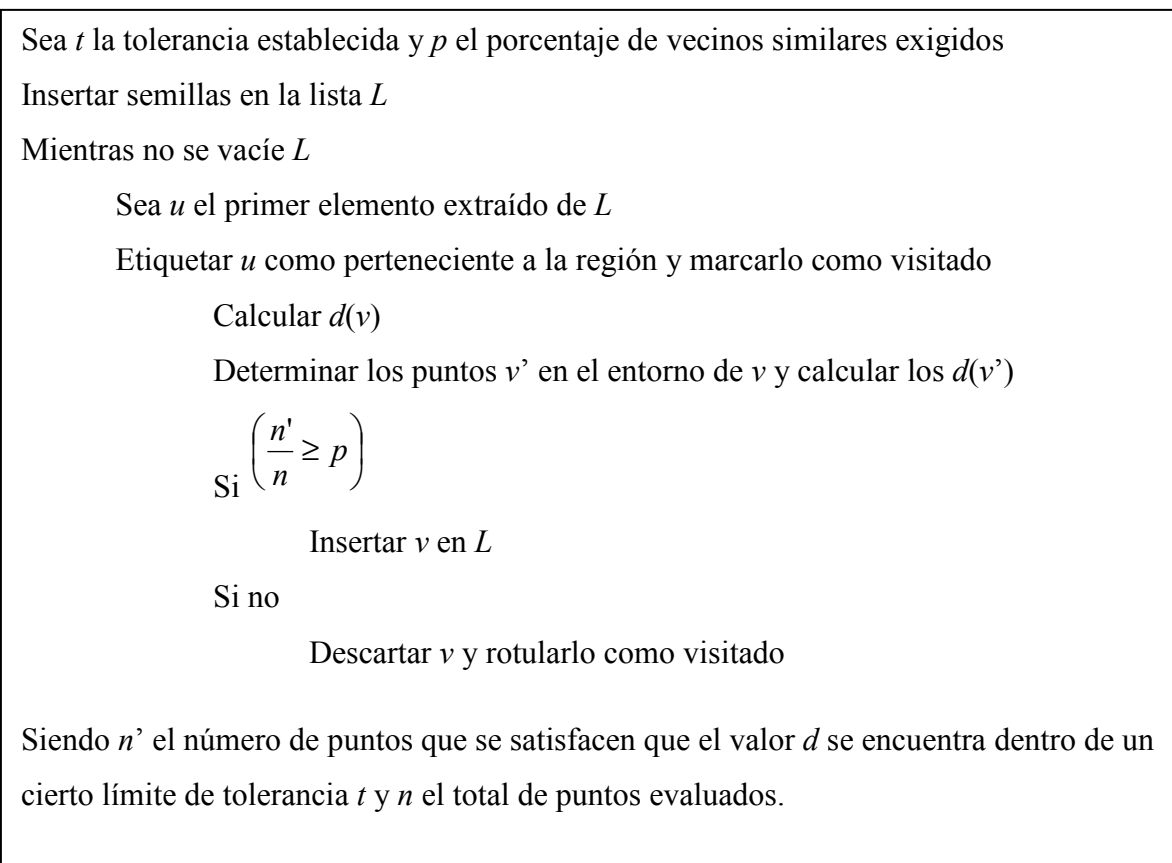


Figura 2.6 Algoritmo de crecimiento de regiones

#### 2.2.2.2. Algoritmo de División y fusión (“Split & Merge”)

Los aspectos en los que se basa el algoritmo de Split & Merge son la división y unión de regiones. Normalmente las dos fases del algoritmo se hacen precisamente en este orden, primero una división en regiones para posteriormente unir aquellas regiones vecinas similares.

Inicialmente la imagen es considerada como una única región cuadrada. Iterativamente, cada una de las regiones que componen la imagen son analizadas comprobando si dicha región es uniforme o no, según algún criterio previamente establecido, tal como que la diferencia de los valores de intensidad de los píxeles no deben diferenciarse entre sí más de un cierto valor respecto de uno central de referencia. Si es uniforme esta región se mantiene intacta. Si no es uniforme, la región es dividida en cuatro partes iguales que serán analizadas en las etapas posteriores. Es el proceso iterativo continúa hasta que las regiones son uniformes o lo suficientemente pequeñas.

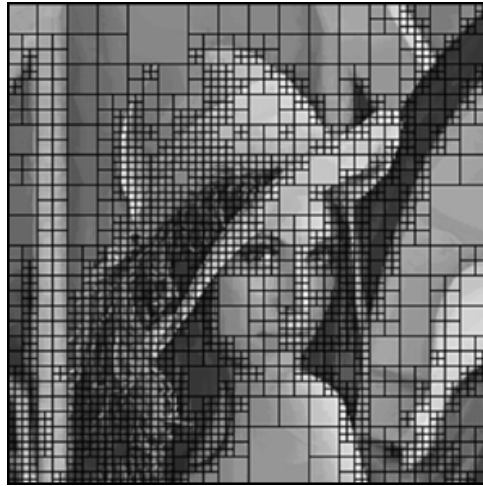


Figura 2.7 Imagen dividida en “quadtrees” mediante el algoritmo de División.

Cuando el proceso de división termina tiene lugar la fase de unión de las regiones obtenidas. En este caso, se ha de comprobar si dos bloques consecutivos, adyacentes, es decir, que estén en contacto espacialmente, cumplen el criterio anterior. Si es así, dichos bloques quedarán fusionados, formando un único bloque. Dicho proceso se repite hasta llegar a la idempotencia, es decir, el momento en el cual ya no pueden fusionarse más bloques.

El criterio adoptado en este proceso es igual que el aplicado en el caso anterior, con el apunte adicional de que el umbral de decisión puede ser variado entre la etapa de división y la de unión. Por lo general se suele tomar un umbral bastante estricto al realizar la división, suavizándolo a continuación en la fase de unión.

### **División**

Se inicializa la lista de imágenes a procesar con la imagen completa.

Repetir hasta que la lista de imágenes a procesar esté vacía

- Extraer el primer elemento de la lista de imágenes a procesar
- Si la región cumple la condición de similitud entonces añadirla a la lista de regiones

Si no dividir la región en cuatro subregiones y añadirla a la lista de imágenes a procesar

Figura 2.8 Algoritmo División

### **Fusión**

Añadir todas las regiones obtenidas en la etapa de división a la lista de imágenes a procesar.

Repetir hasta que no se produzca ninguna unión nueva

- Extraer cada región de la lista de imágenes a procesar y buscar una región vecina. Si al unir ambas regiones se sigue cumpliendo la condición de similitud entonces unir las y añadir la nueva región a la lista.

Figura 2.9 Algoritmo Fusión

#### **2.2.2.3. Watershed**

La técnica conocida como Watershed es una técnica bien conocida en área de la topografía. Esta técnica consiste en interpretar la imagen como una superficie topográfica donde el nivel de gris de un punto indica la altitud en dicho punto. Por tanto, se intenta dividir la imagen de grises en regiones, generalmente una región formará el fondo de la imagen y el resto los objetos a extraer.

La idea básica consiste en considerar que los contornos de los objetos se corresponden con las líneas donde el nivel de gris varía más rápidamente. Para determinar dicha variación, en este trabajo usaremos el gradiente.

La técnica Watershed considera la imagen como un gráfico 3D, en el que se representa la altitud (nivel de gris) frente a las dos coordenadas espaciales de dicho punto. Los mínimos

locales de dicho gráfico representarían cuencas donde el agua se iría acumulando, simulando así el comportamiento del terreno en la realidad. Llegado el punto en el que si seguimos llenando una cuenca ésta se desbordaría, se construye un dique para evitar la fusión de dos o más de estas cuencas. Estos diques serían los contornos de las imágenes a reconocer.

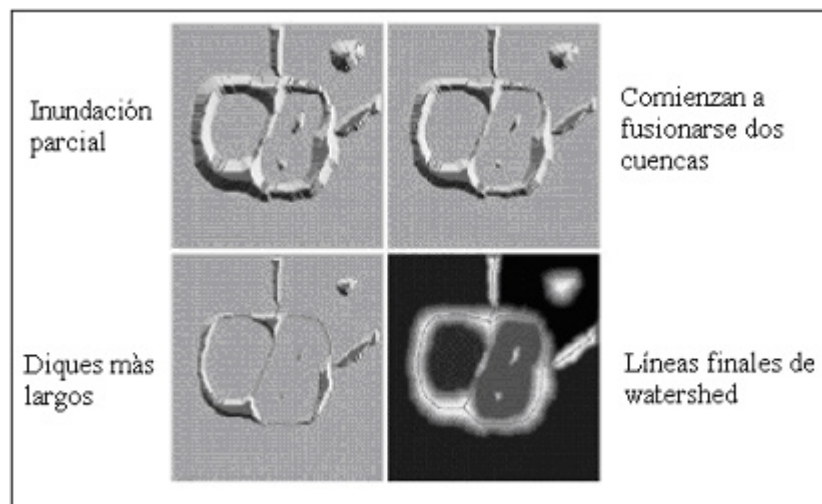


Figura 2.10 Etapas en el algoritmo de watershed

Sin embargo, la técnica de Watershed suele llevar a una sobre-segmentación. Para evitar esto, existen dos posibilidades:

- Eliminar los contornos irrelevantes una vez realizado el proceso de Watershed.
- Modificar la imagen gradiente de tal forma que las regiones de depresión o valles se correspondan únicamente con los objetos deseados.

## 2.3. Visión estéreo

La visión estereoscópica es el proceso por el cual se reconstruye un escenario 3D a partir de dos imágenes 2D. El cerebro humano percibe la realidad tridimensional a través del sistema de visión, que consta de dos ojos desplazados ligeramente uno respecto del otro según la dirección horizontal. Basándonos en este hecho los sistemas de percepción artificiales poseen dos cámaras desplazadas entre sí una cierta distancia, gracias a lo cual se puede recuperar la tercera dimensión perdida en las imágenes bidimensionales.

En la figura 2.11 se muestra un par de imágenes estereoscópicas obtenidas mediante dos cámaras desplazadas entre sí una cierta distancia horizontal.

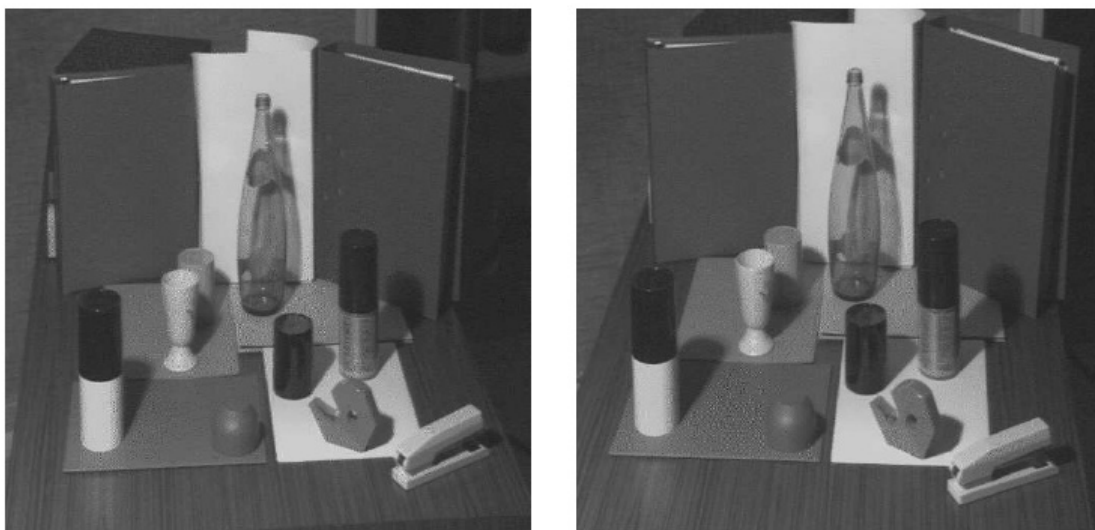


Figura 2.11 Imágenes capturadas por un sistema de visión estéreo

Para estimar la profundidad de dos imágenes estéreo debemos solucionar dos problemas principales:

- Problema de correspondencia: para todos los elementos en la imagen izquierda, encontrar sus elementos correspondientes en la imagen derecha, considerando ‘elementos’ como píxeles, características, regiones, objetos, etc.
- Problema de reconstrucción: usando las disparidades estimadas entre elementos, reconstruir la estructura 3D de la escena. Para esta etapa se necesita información adicional sobre las cámaras y hacer ciertas suposiciones sobre la escena. Esta etapa del proceso no se aborda en este trabajo.

### **2.3.1. Problema de la correspondencia**

Dadas dos o más imágenes del mismo escenario 3D tomadas desde diferentes puntos de vista, el problema de la correspondencia consisten en encontrar conjuntos de puntos en una imagen que puedan ser identificados con los mismos puntos en la otra imagen. Un humano normalmente puede solucionar este problema fácilmente. En visión artificial el problema de la correspondencia se trata de solucionar de forma automática únicamente con las imágenes del par estereoscópico como entrada. Una vez que el problema de la

correspondencia ha sido resuelto, obteniendo un conjunto de puntos de una imagen con sus homólogos en la otra imagen, se pueden aplicar otros métodos para reconstruir la posición de dichos puntos en la escena 3D.

### ***Métodos para establecer la correspondencia***

Existen dos aspectos que deben ser considerados:

- Cómo seleccionar las correspondencias
- Cómo determinar la bondad de las correspondencias

En la literatura aparecen dos clases principales de algoritmos:

- Basados en el área: establecen la correspondencia atendiendo a los niveles de intensidad de la imagen píxel a píxel y teniendo en cuenta todos los píxeles de las imágenes.
- Basados en características: la correspondencia se establece mediante un conjunto de características de la imagen, tales como puntos de borde, puntos de interés, regiones, etc.

#### **2.3.1.1. Métodos basados en el área**

Son métodos que establecen la correspondencia píxel a píxel, utilizando para ello la información, principalmente de intensidad (color) entre un píxel y sus vecinos en una imagen con el píxel y sus vecinos en la otra imagen del par estereoscópico. De este modo, se determina una medida de similitud entre ambos entornos. Uno de los métodos más utilizados para la medida de la similitud consiste en el cómputo del coeficiente de correlación, basado en los mencionados entornos (Pajares y Cruz, 2007).

Para que la correspondencia pueda establecerse correctamente, se requiere que los valores de las intensidades de las imágenes del par estereoscópico sean similares.

En la figura 2.12 se muestran dos imágenes, izquierda y derecha, que constituyen un par estereoscópico. En ellas se han seleccionado una serie de puntos de interés, siendo el objetivo determinar en sendas imágenes cuál es la correspondencia entre ellos. Obsérvese, en este caso la similitud de valores de intensidad entre píxeles homólogos de ambas imágenes.

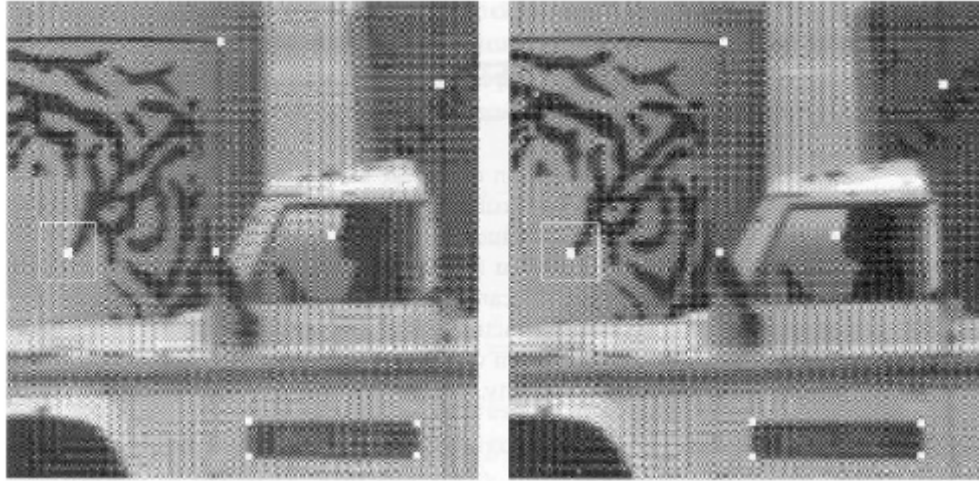


Figura 2.12 Ejemplo de correspondencia de puntos

La estructura básica del algoritmo se describe en la figura 2.13.

Entradas:

- Imagen izquierda ( $I_i$ ) e imagen derecha ( $I_d$ )
- Anchura de la ventana  $N$ , siendo  $N=2W+1$ .
- Región de búsqueda en la imagen derecha  $R(p_i)$  asociado con un píxel  $p_i$  en la imagen izquierda

Para cada píxel  $p_i = (i, j)$  en la imagen izquierda:

- Para cada desplazamiento  $\bar{d} = (\bar{d}_1, \bar{d}_2) \in R(p_i)$  calcular la similitud entre la ventana en la imagen izquierda y la ventana en la imagen derecha ( $c(d)$ )
- La disparidad de  $p_i$  es el vector  $\bar{d} = (\bar{d}_1, \bar{d}_2)$  que maximiza  $c(d)$  sobre  $R(p_i)$

$$d = \arg \max_{d \in R} [c(d)]$$

Figura 2.13 Algoritmo de correspondencia basada en el área.

Considerando ventanas de  $N \times N$  centradas en los puntos a estudiar y donde  $I_i(i, j)$  e  $I_d(i, j)$  denotan las intensidades de un píxel en la posición  $(i, j)$  pueden considerarse las siguientes funciones como medidas de similitud:

- Correlación cruzada (Anuta, 1969):

$$c(d) = \sum_{k=-W}^W \sum_{l=-W}^W I_i(i+k, j+l) I_d(i+k-d_1, j+l-d_2) \quad (2.5)$$

Normalmente dividiremos (2.5) entre la desviación estándar de  $I_i$  y  $I_d$ :

$$\bar{c}(d) = \frac{\sum_{k=-W}^W \sum_{l=-W}^W (I_i(i+k, j+l) - \bar{I}_i)(I_d(i+k-d_1, j+l-d_2) - \bar{I}_d)}{\sqrt{\sum_{k=-W}^W \sum_{l=-W}^W (I_i(i+k, j+l) - \bar{I}_i)^2 \sum_{k=-W}^W \sum_{l=-W}^W (I_d(i+k-d_1, j+l-d_2) - \bar{I}_d)^2}} \quad (2.6)$$

- Suma de diferencias absolutas (Barnea y Silverman, 1972):

$$\bar{c}(d) = - \sum_{k=-W}^W \sum_{l=-W}^W |I_i(i+k, j+l) - I_d(i+k-d_1, j+l-d_2)| \quad (2.7)$$

- Suma de los cuadrados de las diferencias (Shirai, 1987):

$$\bar{c}(d) = - \sum_{k=-W}^W \sum_{l=-W}^W (I_i(i+k, j+l) - I_d(i+k-d_1, j+l-d_2))^2 \quad (2.8)$$

El éxito de los métodos basados en correlación depende de que la ventana utilizada capture la estructura que rodea al punto que se está estudiando para establecer la correspondencia.

### *¿Cómo elegir el tamaño de la ventana?*

Una ventana demasiado pequeña puede no capturar suficiente estructura de la imagen, y puede ser demasiado sensible a ruido. Una ventana demasiado grande es menos sensible al ruido, pero a cambio está sujeta a posibles variaciones de intensidad entre las dos imágenes del par estereoscópico.

Realizar la correlación sobre todos y cada uno de los píxeles de la imagen homóloga requiere un elevado tiempo de computación. Para mejorar el tiempo de respuesta de la correlación añadimos la restricción epipolar, que consiste en limitar la búsqueda sobre el epipolo de la imagen homóloga. El epipolo es el resultado de la proyección en la cámara homóloga de la línea de proyección que causó el píxel a emparejar.

#### **2.3.1.2. Métodos basados en las características**

Estos métodos tratan de emparejar una característica en una imagen con otra característica de propiedades similares en la otra imagen. En la literatura se han utilizado diversas características, entre las cuales destacan:



- puntos de borde. (Grimson, 1985; Kim y Aggarwal, 1987; Khotanzad y col., 1993; Lew y col., 1994; Lutsiv y Novikova, 1992; Maravall y Fernandez, 1992; Marr, 1982, 1985; Marr y Poggio, 1976, 1979; Mousavi y Shalkoff, 1994; Rubio, 1993; o Ruycheck y Postaire, 1996).
- segmentos de línea. Medioni y Nevatia (1985, 1984) que aplican el principio de Disparidad Diferencial Mínima; Ayache y Faverjon (1987) que emplean el Límite del Gradiente de Disparidad; Pajares y col. (1993) y Kim y col. (1992) que aplican técnicas de Lógica Difusa; Nasrabadi (1992) y Li (1994) que utilizan segmentos curvos y aplican la transformada de Hough; Boyer y Kak (1988) emplean relaciones geométricas entre las estructuras que conforman la imagen. Ohta y Kanade (1985), Shirai (1987) o Lee y Leou (1994) usan técnicas de programación dinámica; Krotkov y col. (1990) trabajan con sistemas convergentes; Tanaka y Kak (1990) introducen un cierto conocimiento previo de superficies; Tu y Dubuisson (1990) se basan fundamentalmente en la restricción epipolar. Cruz y col. (1995a y b), Pajares y col. (1999, 1998 b-d, 1997 a-c), Pajares y col. (1998a, b, c y d) y Pajares y col. (1997a, b y c) utilizan diversas técnicas de reconocimiento de patrones y aprendizaje. Pajares y col. (1999 c) emplean un método de relajación. Pajares y col. (1998a) que emplea un procedimiento de optimización basado en el modelo de red neuronal de Hopfield.
- esquinas. Wang y col (1983).

En la figura 2.14 vemos un posible ejemplo de correspondencia basada en características en la que se van a comparar dos segmentados de línea (L1 y L2), dos esquinas (E1 y E2) y los puntos de borde de dos regiones (B1 y B2).

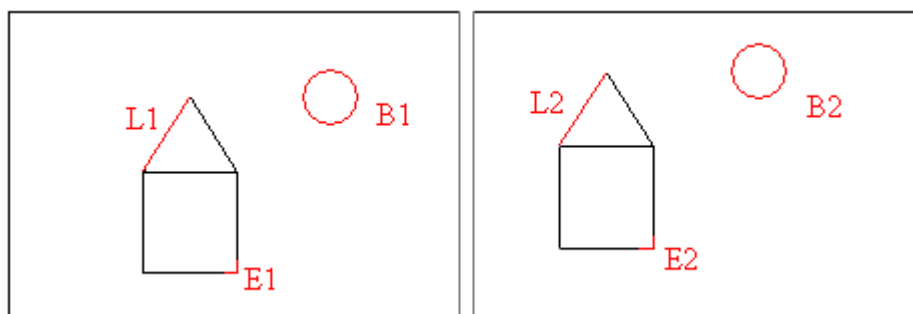


Figura 2.14 Imágenes con características capturadas

Una característica se describe mediante una serie de propiedades, que van a permitir establecer las correspondencias por similitud entre los valores de las propiedades de una característica en una imagen y las características en la otra imagen del par estereoscópico. A modo de ejemplo, un descriptor de línea podría describirse mediante las siguientes propiedades:

- la longitud,  $l$ ;
- la orientación,  $\theta$ ;
- las coordenadas del punto medio,  $m$ ;
- la media de intensidad en la línea,  $i$ .

Cada una de las propiedades mencionadas posee un peso específico concreto, de forma que su importancia relativa quede fijada por el valor de dicho peso a la hora de establecer la correspondencia. Suponiendo  $w_0, \dots, w_3$  los pesos asignados, la similitud entre la imagen izquierda y la imagen derecha viene dada por la ecuación (2.9)

$$S = \frac{1}{w_0(I_i - I_d)^2 + w_1(\theta_i - \theta_d)^2 + w_2(m_i - m_d)^2 + w_3(i_i - i_d)^2} \quad (2.9)$$

El algoritmo general para la correspondencia por características es el mostrado en la figura 2.15.

Entradas:

- Imagen izquierda ( $I_i$ ) e imagen derecha ( $I_d$ )
- características y descriptores de ambas imágenes
- Región de búsqueda en la imagen derecha  $R(f_i)$  asociada con una característica  $f_i$  en la imagen izquierda

Para cada característica  $f_i$  en la imagen izquierda:

- Calcular la similitud entre  $f_i$  y cada característica de imagen en  $R(f_i)$
- Seleccionar la característica en la imagen derecha  $f_d$ , que maximiza la medida de similitud.
- Guardar la correspondencia y disparidad  $d(f_i, f_d)$

Figura 2.15 Algoritmo de correspondencia basada en características.

### **2.3.1.3. Métodos basados en el área frente a métodos basados en características**

En Scharstein y Szeliski (2001), podemos ver una comparación de distintos métodos de correspondencia. Tanto los métodos basados en el área como los basados en las características, tienen sus ventajas e inconvenientes. A continuación se enumeran algunas de ellas:

- *Métodos basados en el área*
  - Más fáciles de implementar que los métodos basados en características
  - Ofrece un mapa de disparidad denso (útil para reconstruir superficies).
  - Necesitan imágenes con texturas para funcionar correctamente, de lo contrario produce muchas falsas correspondencias.
  - No funcionan bien con puntos de vista muy diferentes (debido a los cambios producidos por la iluminación).
- *Métodos basados en características:*
  - Útil cuando se pueden extraer buenas características de la escena.
  - Más rápidos que los métodos basados en correlación.

- Ofrece un mapa de disparidad escaso (bueno para aplicaciones de navegación)
- Relativamente insensible a cambios de iluminación.

### 3. Clasificación de texturas y correspondencia estereoscópica

#### 3.1. Análisis de las imágenes: problemática

Este capítulo aborda la problemática estereoscópica para un tipo de sistemas característicos como son los formados por dos cámaras cuyas imágenes se obtienen mediante un tipo de lentes conocidas como de “ojo de pez”. En ellas el entorno tridimensional se proyecta de forma que se genera una imagen con proyección de los 360° alrededor de las cámaras. En las figuras 3.1 y 3.2 se muestran ejemplos representativos de imágenes capturadas con estos sistemas.

Estos sistemas se caracterizan, además porque las cámaras se sitúan entre ellas separadas una cierta distancia, en el caso que nos ocupa de un metro. Esto hace que un mismo objeto se proyecte en cada imagen en posiciones distintas. Obsérvese a este respecto en la 3.1 cualquiera de los troncos de los árboles, un mismo tronco aparece en una imagen en una determinada posición y en la otra imagen del par aparece girado un cierto ángulo con respecto al primero. Este hecho es el que permite determinar la estructura tridimensional de la escena, de suerte que es posible establecer la distancia a las que se encuentra el tronco en cuestión respecto del sistema de referencia de las cámaras. El objetivo del trabajo se centra precisamente, en determinar ese desplazamiento relativo para un mismo tronco entre las dos imágenes del par estereoscópico. Dicho desplazamiento se mide en grados. Dada la complejidad de este tipo de imágenes, como se verá posteriormente, el problema se desglosa en dos partes. En primer lugar, como el interés se centra exclusivamente en determinar las posiciones relativas de los troncos, es evidente que el resto de elementos presentes en la escena carecen de interés, este es el caso de las ramas de los árboles, el suelo o el cielo. Por ello, se enfoca el problema tratando de identificar aquellas texturas que son o no relevantes para el problema propuesto. Se trata de abordar esta problemática

mediante técnicas de segmentación e identificación de texturas como las descritas en el capítulo dos.

Una vez identificadas las texturas de interés y descartadas las que carecen de ello, el siguiente paso consiste en establecer las correspondencias entre los píxeles que constituyen los troncos de los árboles. Esta parte es la que coincide con el problema de la correspondencia en visión estereoscópica que se ha descrito en el capítulo dos.

Por tanto, en este capítulo se describen las técnicas utilizadas, en algunos casos de forma independiente y en otros casos de forma combinada para mejorar los resultados obtenidos desde la perspectiva de segmentación de texturas y desde la correspondencia, siguiendo los métodos descritos en el capítulo dos.

Como ya ha sido comentado, el objetivo final del algoritmo consiste en aislar dentro de las imágenes exactamente los troncos de los árboles para poder extraer las propiedades anteriores con las que establecer la correspondencia en un sistema estéreo. Una vez establecida la correspondencia debe ser posible el cálculo del volumen de madera existente en las imágenes, cuestión ésta que no se aborda en el presente trabajo.

A la vista de las imágenes analizadas se identifica la problemática que a continuación se describe según los tipos de imágenes mostrados:

### ***Imágenes captadas con escasa iluminación***

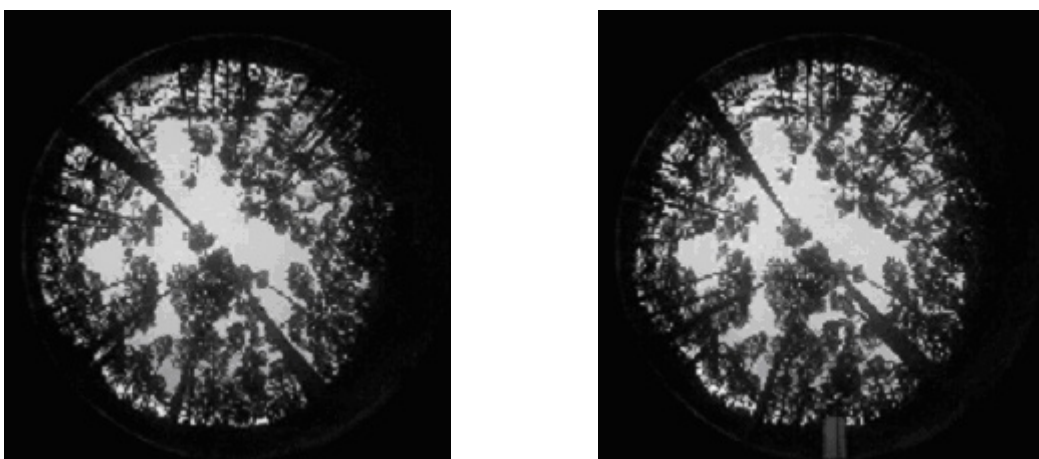


Figura 3.1 Par de imágenes estereoscópicas

### ***Color:***

En este caso los troncos de los árboles presentan desde la base al menos tres tonalidades de color comenzando desde las negras y evolucionando hacia las copas de los árboles donde alcanza diferentes niveles o tonalidades (que resultan ser prácticamente grises). Esto implica que un mismo tronco de árbol no posee la misma textura de color debido a los diferentes niveles de iluminación.

Las tonalidades de los troncos son similares a las del resto de texturas existentes según la situación de la vegetación en la imagen.

Conclusión: no se pueden identificar los troncos sólo por la textura de color debido a su similitud con la maleza, cosa que no ocurre con las texturas del cielo que sí admiten una fácil discriminación

### ***Descriptores de textura estadísticos:***

En este tipo de imágenes es posible distinguir a través de los momentos estadísticos de segundo orden (varianza) diferentes texturas, diferenciándose las texturas de las ramas de las de los troncos por presentar en este último caso una cierta variabilidad, mientras que en el primero se manifiestan de forma homogénea.

### ***Imágenes con variedad de color***

En este tipo de imágenes resulta ya más fácil separar la parte predominantemente verde de la que no lo es. Esta última serían los troncos. Si bien, éstos presentan una problemática que es su distinta tonalidad; en efecto, se pueden distinguir varias tonalidades:

Oscuras: quedan en la parte de la sombra

Claros: por recibir la iluminación solar directamente, en este caso llega a tener el mismo valor que parte del cielo, con lo que la problemática respecto de su posible clasificación se incrementa considerablemente.

Estas dos apreciaciones se manifiestan en la imagen de la izquierda y no así en la imagen de la derecha, donde dado el espesor de vegetación no existen sombras y sólo cabe

distinguir entre verde, azul, blanco y oscuro. Siendo la parte oscura la correspondiente a los troncos.

Se aprecian también diferentes texturas. Distinguiéndose entre las texturas en las zonas donde hay hojas y las texturas de los troncos. Las primeras presentan alta variabilidad entremezclándose verde con azul, mientras que las texturas que representan los troncos no manifiestan este comportamiento.

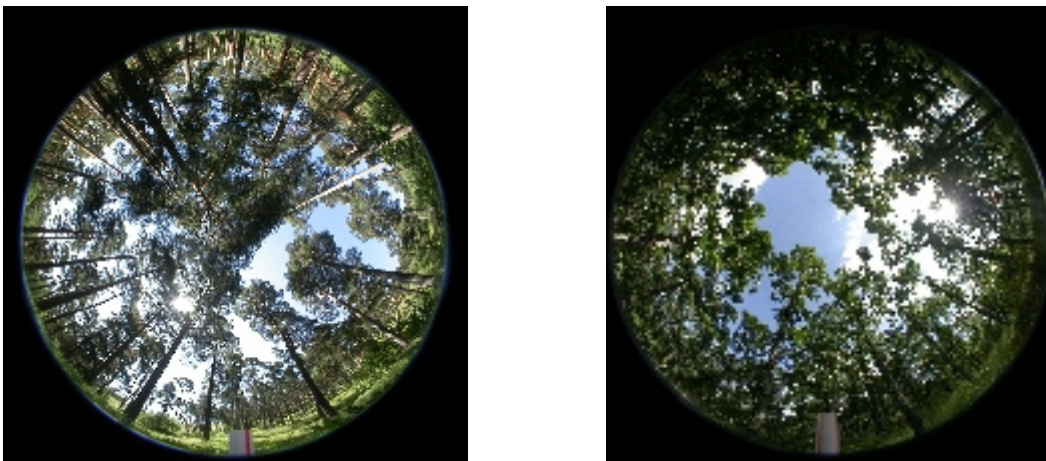


Figura 3.2 Imágenes captadas con sistema estéreo

## 3.2. Segmentación

### 3.2.1. Segmentación basada en la umbralización por histograma

Como se ha mencionado en el capítulo dos, la segmentación basada en la umbralización por histograma aprovecha la similitud de niveles de intensidad en una imagen para establecer si un píxel pertenece a la clase buscada. Este método funciona bien cuando los objetos buscados tienen un nivel de intensidad concreto que lo distingue del resto de objetos.

En la figura 3.3 podemos ver una imagen forestal y su histograma.



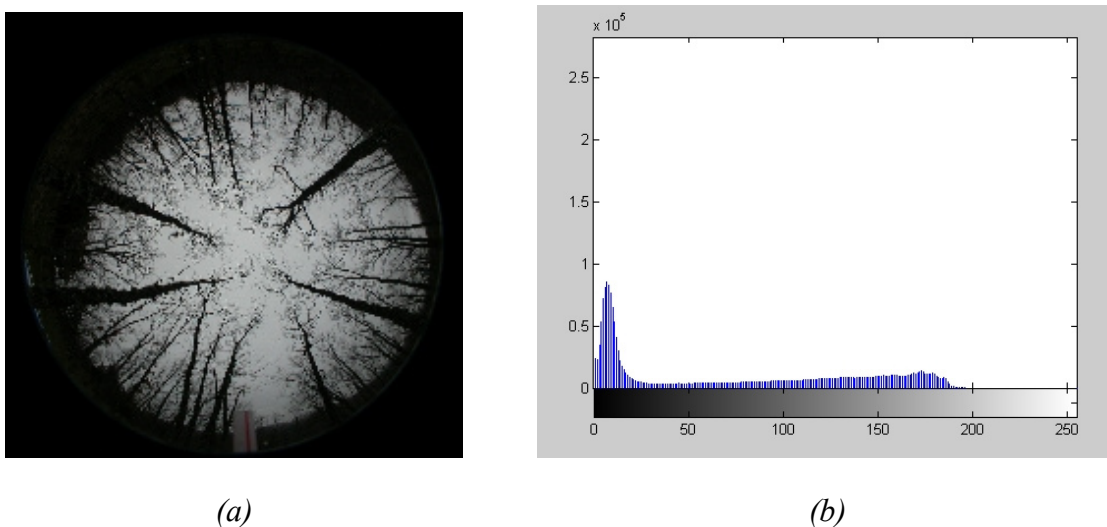


Figura 3.3 (a) Imagen forestal. (b) Histograma de (a).

A la vista del histograma de la figura 3.3 podemos encontrar fácilmente las dos clases que subyacen en la imagen. A la primera clase pertenecen niveles de intensidad entre 0 y 25 y a la segunda clase pertenecen niveles de intensidad entre 26 y 255. Aplicando esta división en dos clases obtenemos la imagen binaria de la figura 3.4.

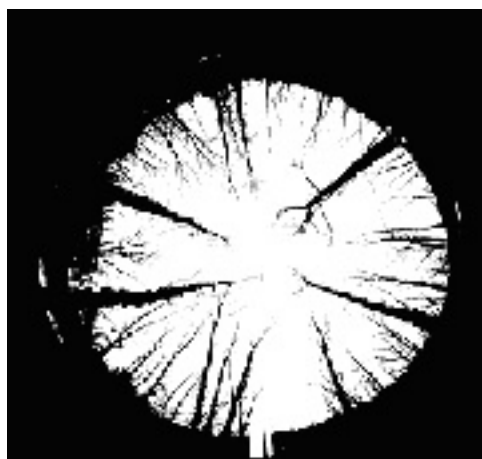


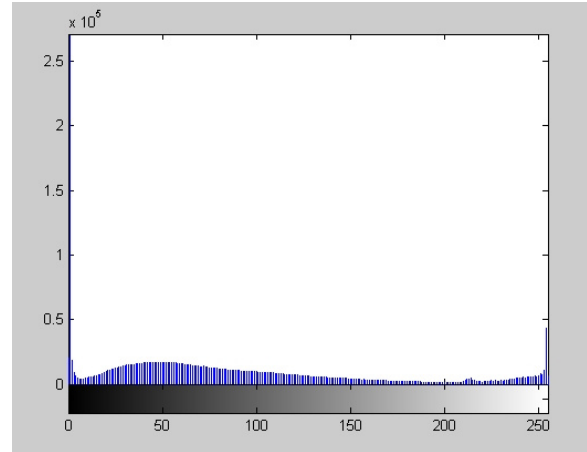
Figura 3.4 Umbralización por histograma de la figura 3.3.

Esta división ha conseguido separar el cielo de los troncos. La siguiente fase consiste en recortar la parte de imagen perteneciente al suelo.

No siempre las imágenes son tan sencillas como la anterior. En la figura 3.5 se muestra otra de las imágenes forestales capturadas con su correspondiente histograma.



(a)



(b)

Figura 3.5 (a) Imagen de color compleja. (b) Histograma de (a).

En esta imagen las clases ya no aparecen tan diferenciadas como en el caso de las imágenes anteriores. Un posible umbral es 75. Usando este umbral se obtiene la imagen binaria que aparece en la figura 2.6

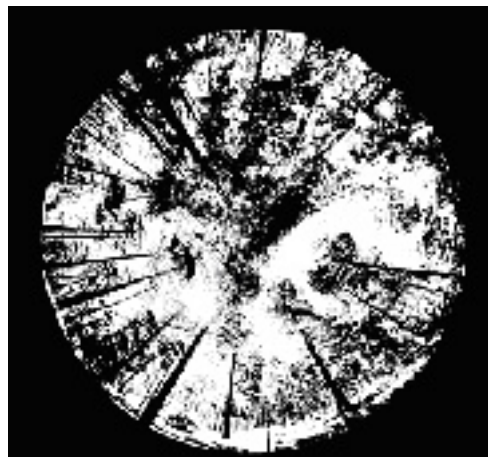


Figura 3.6 Umbralización de la figura 3.5(a).

La segmentación realizada ha conseguido capturar los troncos pero junto a los troncos han quedado partes de hojas. Esto se debe a que ambas clases de texturas poseen niveles de intensidad muy semejantes.

Si usamos el método de Otsu (1979) para calcular el valor umbral se obtiene un umbral de 94.98. En la figura 3.7 se muestra la imagen binaria obtenida tras la umbralización con este valor.

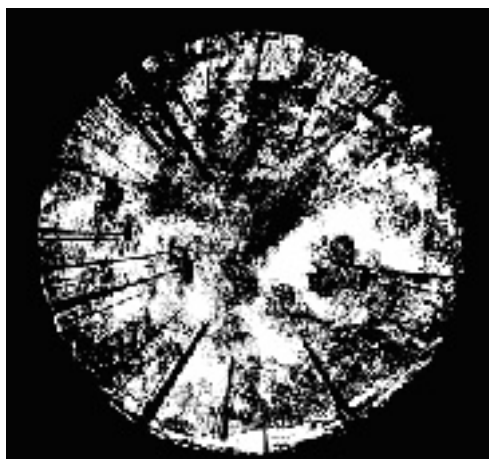


Figura 3.7 Umbralización de la figura 3.5 con método de Otsu.

Como podemos ver, el problema mencionado anteriormente persiste. Queda claro que en esta imagen no es suficiente con una umbralización por histograma. Será necesario aplicar métodos adicionales para eliminar los objetos que no pertenecen a troncos de árbol.

### **3.2.2. Segmentación basada en regiones**

Como se ha comentado anteriormente, en este tipo de segmentación será necesario definir cómo iniciar la partición y definir un criterio de homogeneidad para determinar qué regiones son aceptadas como troncos y cuáles no.

A continuación se indican las técnicas con las que se ha experimentado.

#### **3.2.2.1. Algoritmo de crecimiento de regiones**

Los puntos clave en este algoritmo son la selección de semillas y el criterio de aceptación de puntos dentro de una región, tal y como quedó patente en la exposición realizada en el capítulo dos. Como ya se definió allí, una semilla es un punto a partir del cual crece una estructura que identifica la región. Para la selección de semillas se han usado dos técnicas, una supervisada y otra no supervisada:

- Selección de semillas por un experto: en este caso, un experto se encarga de seleccionar los puntos clave a partir de los que se deben expandir las regiones. Estos puntos deben coincidir con un punto de cada tronco.
- Utilización de cada píxel como posible semilla: para cada punto en la imagen que no ha sido etiquetado dentro de una región se expande una nueva región a partir de él.

La primera opción tiene la desventaja de no ser un proceso totalmente automático. Sin embargo presenta un mejor rendimiento por no necesitar un recorrido exhaustivo de la imagen. El segundo método necesita además alguna técnica adicional que permita conocer que región calculada pertenece a un tronco y cuál no.

El segundo aspecto a tratar es el criterio de aceptación de un píxel dentro de una región. Para decidir si un punto pertenece a una región o no, sólo se han considerado los niveles de intensidad de dicho píxel, considerar también los vecinos de este píxel conlleva tiempos de cálculos demasiado elevados. El nivel de intensidad de dicho píxel puede compararse con el nivel de intensidad de la semilla solamente, con el nivel de intensidad del píxel vecino a partir del que se ha llegado o a partir del nivel de intensidad medio de la región.

### ***Método supervisado de selección de semillas***

En la figura 3.8 se muestran las regiones obtenidas utilizando la selección de semillas con el método supervisado para algunas de las imágenes capturadas. Podemos ver la semilla seleccionada y el resultado de realizar un crecimiento de región a partir de dicha semilla seleccionada de forma manual.

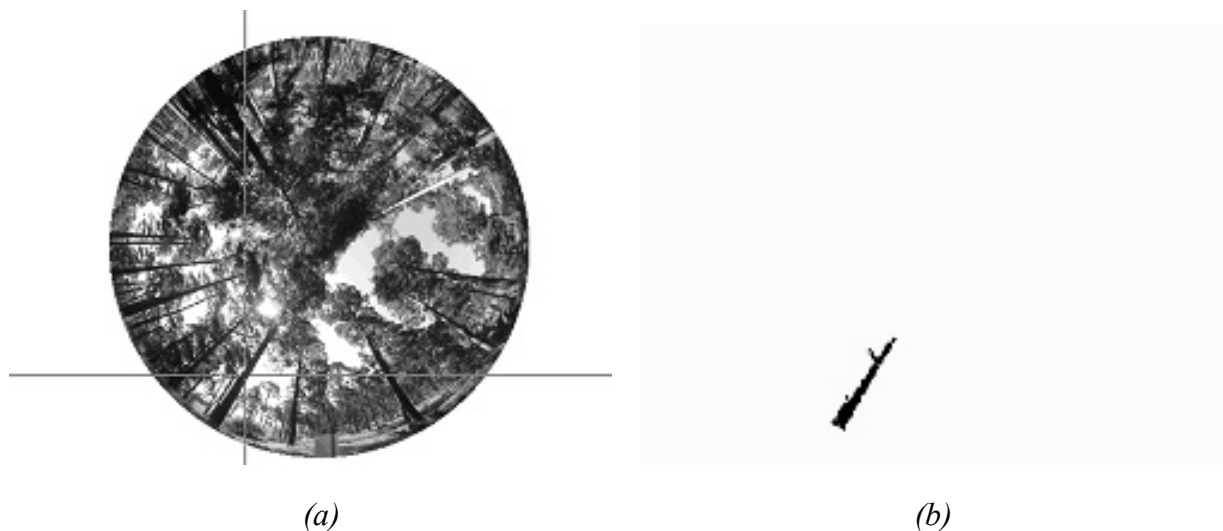


Figura 3.8 Segmentación supervisada. (a) Selección de semilla. (b) Región segmentada sobre región homogénea.

El problema que presenta esta técnica es que sólo funciona bien cuando toda la región del tronco es homogénea. Si el tronco presenta varias regiones, sólo se obtendrá la región a la que pertenezca la semilla seleccionada. En la figura 3.9 podemos observar el problema

mencionado para el tronco en cuestión que presenta diferentes tipos de iluminación y por tanto de texturas.

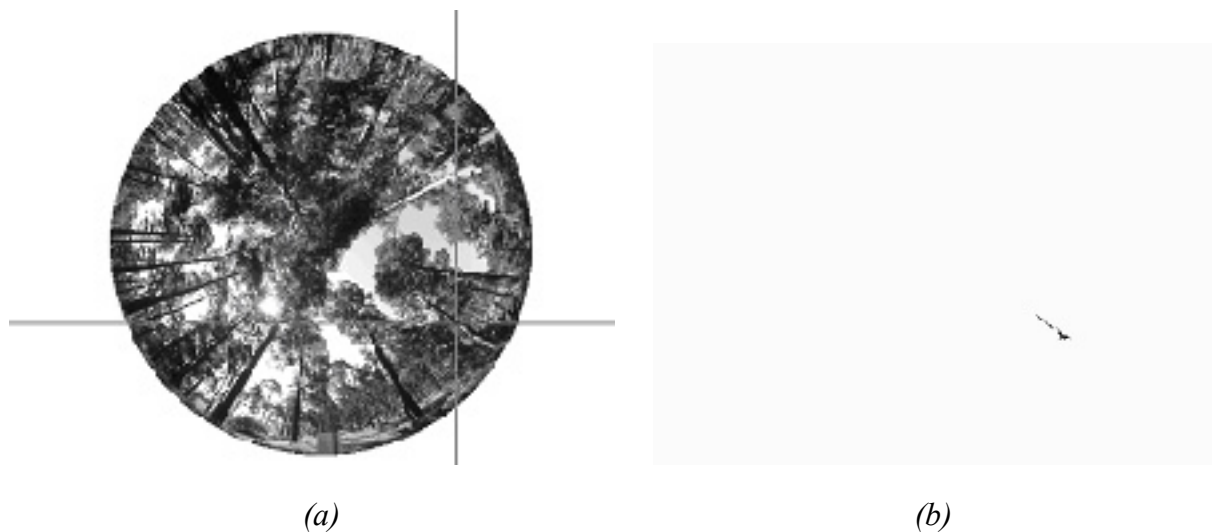


Figura 3.9 Segmentación supervisada. (a) Selección de semilla. (b) Región segmentada sobre región no homogénea.

Una posible solución a este problema es obligar al experto a seleccionar todas y cada una de las subregiones pertenecientes a un tronco. Este proceso puede ser excesivamente complejo y tedioso por lo que se ha optado por buscar un método automático de selección de regiones.

#### ***Método no supervisado de selección de semillas***

Como ya se ha descrito, el proceso automático para buscar las regiones deseadas debe separar toda la imagen en regiones para después seleccionar aquéllas que pertenecen a troncos de árboles. Para ello se recorre cada píxel en la imagen, si este píxel ha sido ya etiquetado en una región no se considera. Por el contrario si no ha sido etiquetado aún entonces se crea una nueva región utilizando como semilla este píxel. En la figura 3.10 podemos ver el resultado obtenido, (b), para una imagen forestal (a). Un hecho destacable es que dicho resultado depende de la tolerancia permitida al aceptar un píxel como perteneciente a una región en función de la similitud del color utilizada.

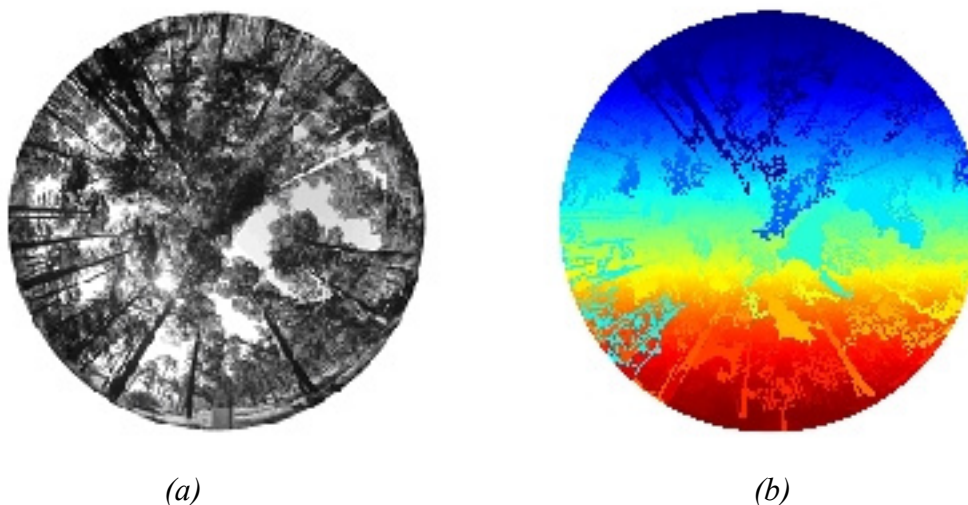


Figura 3.10 Segmentación no supervisada de regiones. (a) Imagen capturada. (b) Regiones extraídas en la segmentación.

Como puede observarse en la figura, algunas regiones han quedado bien separadas, sin embargo otras han quedado fusionadas con regiones formadas por hojas. El problema que presenta la segmentación de regiones forestales consiste en encontrar un umbral que nos permita discernir sobre si un píxel pertenece a una región u otra. Tras obtener estas regiones es necesario un criterio de aceptación de regiones como regiones propias de troncos. Entre los criterios que podemos usar para saber si una región pertenece a un tronco están:

- Nivel medio de intensidad que presenta la región.
- Forma de las regiones: los troncos deben ser regiones alargadas.
- Orientación: las regiones pertenecientes a troncos deben estar orientadas hacia el centro de la imagen.
- Área: puede usarse un área mínima como restricción para clasificar a una región como tronco o no.

Estas técnicas sólo se han intuido, sin haber llegado a aplicarlas al problema de identificación de las texturas, por lo que su estudio se pospone para futuros trabajos de investigación.

### 3.2.2.2. Algoritmo de División y Fusión “Split&Merge”

El algoritmo de crecimiento de regiones es excesivamente sensible a variaciones de iluminación en determinados píxeles. El algoritmo *división y fusión* (Split&Merge) considera texturas en vez de píxeles a nivel individual por lo que es menos sensible a cambios en los valores de intensidad o color a nivel de píxel.

A la hora de diseñar el algoritmo de división debemos decidir el tamaño mínimo de ventana que se considera para tratar las texturas junto con el criterio de homogeneidad de la región para llevar a cabo la división. Como criterio de homogeneidad se ha usado la desviación estándar y la media de color de la textura. Si la desviación no supera un cierto valor y la media está en un rango de valores se acepta la textura, en caso contrario se rechaza. En nuestro caso se han aceptado regiones con una desviación estándar no superior a 5 y una media de intensidad no superior a 150. De esta forma se eliminan las regiones con hojas y las regiones del cielo.

Se han realizado varias pruebas con distintos tamaños de ventana ( $N$ ). En la figura 3.11 se muestran los resultados obtenidos usando división y fusión con tamaños de ventana 3, 5, 7 y 11 sobre 3.11(a).



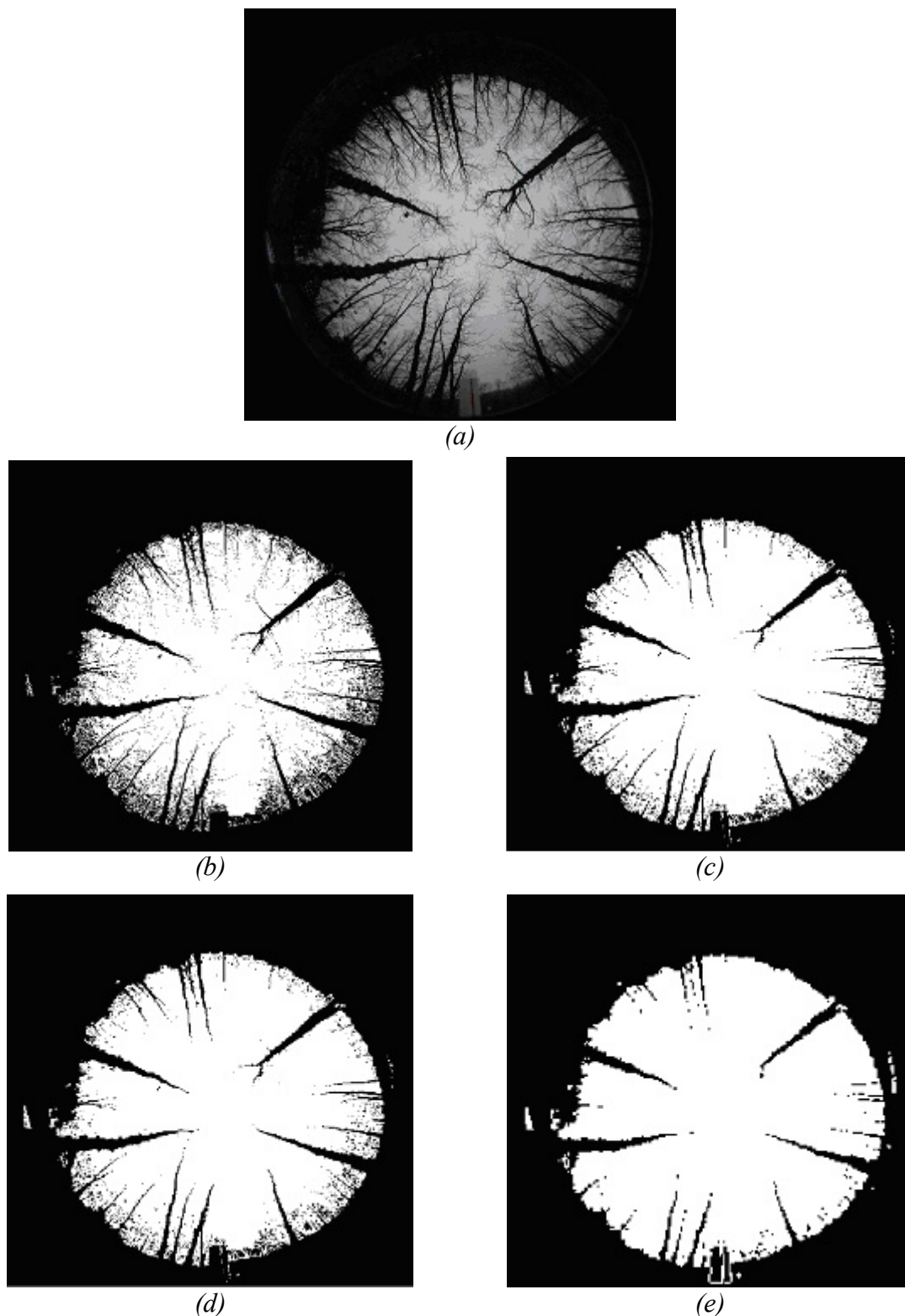


Figura 3.11 Algoritmo de Split&Merge usando distintos tamaños de ventana ( $N$ ). (a) Imagen original. (b) Segmentación con  $N=3$ . (c) Segmentación con  $N=5$ . (d) Segmentación con  $N=7$ . (e) Segmentación con  $N=11$ .

Como se puede observar en la 3.11(b), un tamaño de ventana demasiado pequeño produce sobre-segmentación, mientras que con un tamaño de ventana grande podemos perder información como puede verse en 3.11(e).



A partir de la imagen con tamaño de ventana 11 podemos eliminar las regiones con un área pequeña obteniendo la imagen mostrada en la figura 3.12.

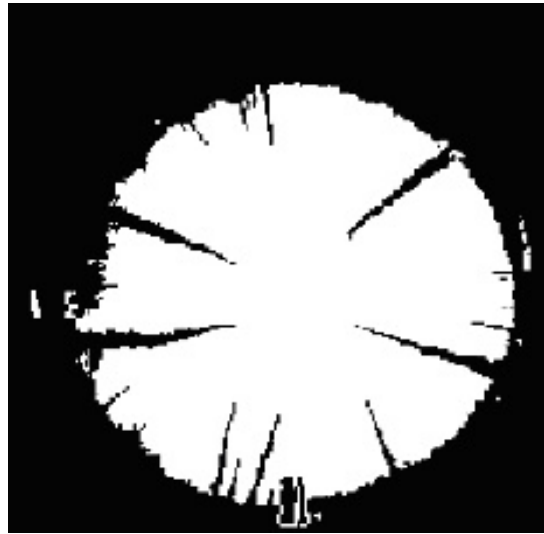


Figura .3.12 Segmentación eliminando regiones pequeñas

A partir del resultado obtenido en la figura 3.12 y eliminando la parte externa de la circunferencia y las áreas resultantes con baja excentricidad llegamos al resultado de la figura 3.13.

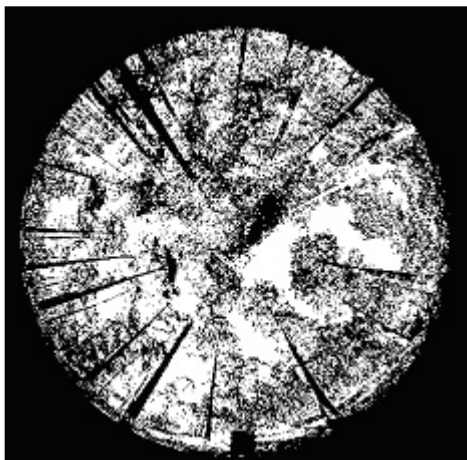


Figura 3.13 Segmentación eliminando suelo

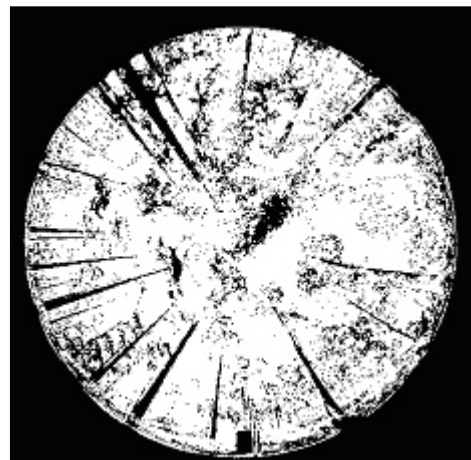
Tal y como se puede apreciar, el resultado obtenido sobre esta imagen resulta ser satisfactorio mediante la aplicación del método de segmentación expuesto. La dificultad se presenta cuando tenemos imágenes más complejas como la mostrada en la figura 3.14.



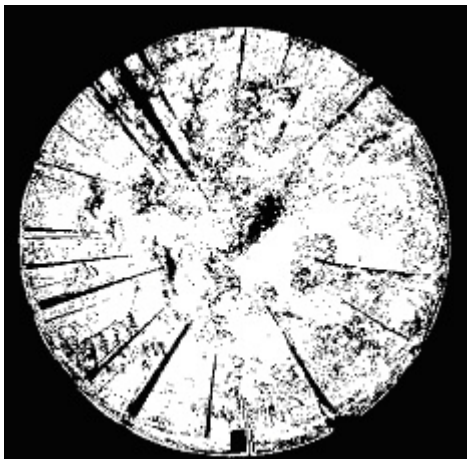
(a)



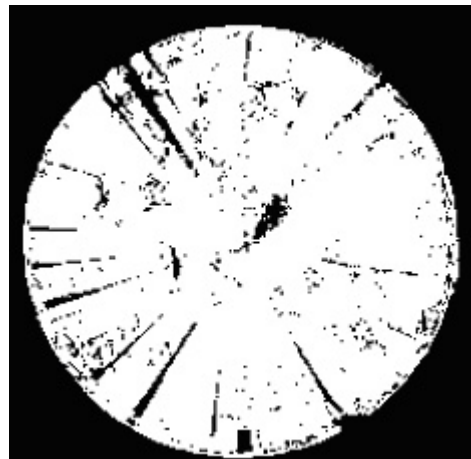
(b)



(c)



(d)



(e)

Figura 3.14 Algoritmo de Split&Merge sobre imagen compleja usando distintos tamaños de ventana (N). (a) Imagen original. (b) Segmentación con  $N=3$ . (c) Segmentación con  $N=5$ . (d) Segmentación con  $N=7$ . (e) Segmentación con  $N=11$ .

Podemos observar que en las imágenes segmentadas han quedado regiones que no son troncos y algunos troncos han desaparecido.

Si nos quedamos con la imagen 3.14(e), podemos mejorar la imagen eliminando regiones pequeñas y regiones con excentricidad alta. Esta operación se realiza mediante la extracción de propiedades de las regiones tales como el área y la excentricidad de las mismas. Los experimentos realizados en este trabajo se han llevado a cabo con Matlab, que posee la función *regionprops*, con la que podemos obtener las propiedades mencionadas de área y excentricidad de cada una de las regiones.



Figura 3.15 Segmentación eliminando regiones pequeñas

A pesar de haber conseguido una mejora en cuanto a los resultados de segmentación, todavía no se ha conseguido llegar a resultados satisfactorios, ya que se mantienen regiones que no son troncos de árboles y se han perdido otra serie de árboles que deberían haber permanecido.

Aplicando un cierto conocimiento de índole superior, se deduce que aquellas regiones pertenecientes a un tronco de árbol deben estar cerca el suelo por lo que podemos eliminar aquellas regiones que no estén cerca del suelo, obteniendo en este caso el resultado que se muestra en la figura 3.16.



Figura 3.16 Segmentación eliminando regiones lejos del suelo

A la vista de los resultados obtenidos con este método de segmentación, podemos deducir que este algoritmo por sí mismo tampoco resulta ser satisfactorio para el proceso de segmentación que se pretende.

### 3.2.2.3. Watershed

Como ya se comentó en el capítulo dos, el algoritmo de Watershed produce sobre-segmentación en imágenes complejas. Para imágenes sencillas como la de la figura 3.3 se obtienen buenos resultados debido a que las regiones están bien diferenciadas. Para imágenes más complejas como la de la figura 3.4 se produce sobre-segmentación como puede apreciarse en los resultados mostrados en la figura 3.17.

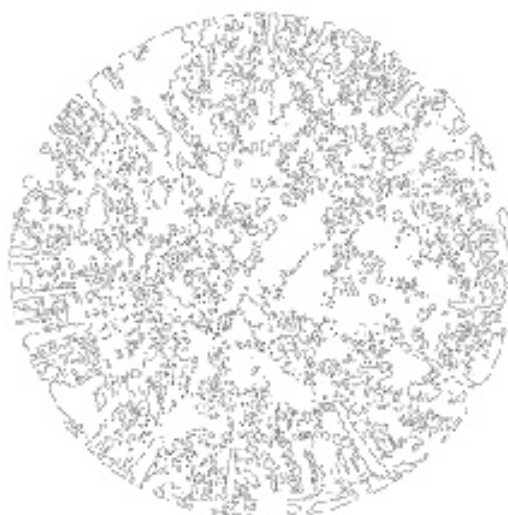


Figura 3.17 Segmentación usando watershed

Se ha implementado el algoritmo de Watershed con distintas modificaciones para evitar este problema sin conseguir resultados apreciables. Este método, a la vista de los experimentos realizados, no parece aportar mejoras respecto a los algoritmos anteriores.

#### **3.2.2.4. Resumen de los métodos de segmentación**

A la vista de los resultados mostrados previamente, no se puede afirmar la existencia de un método válido con carácter universal para la segmentación de las texturas en este tipo de imágenes. Como consecuencia de ello, lo único que se puede concluir es la falta de eficacia de los métodos de clasificación de texturas para el mencionado tipo de imágenes.

### **3.3. Correspondencia**

Con el problema de la correspondencia se trata de identificar cada píxel de la imagen izquierda en la imagen derecha.

En la figura 3.18 se muestra la correspondencia de un punto en la imagen izquierda con su homólogo en la imagen derecha. Como las cámaras que forman el sistema estéreo tienen perspectivas distintas, el mismo punto se proyecta en localizaciones diferentes.

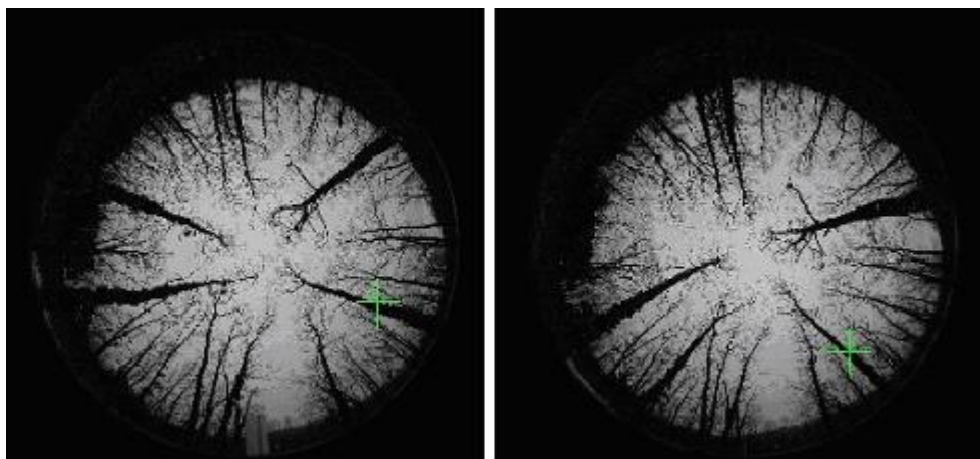


Figura 3.18 Problema de la correspondencia

Como ya se ha mencionado anteriormente, la correspondencia de imágenes es un proceso mediante el cual se trata de identificar la misma característica espacial en ambas imágenes del par estereoscópico. Desde el punto de vista humano, el problema resulta relativamente sencillo, ya que fijándose en un punto de una imagen resulta relativamente fácil identificar su homólogo en la otra imagen del par estereoscópico. Sin embargo, desde el punto de

vista de la visión por computador el proceso se complica enormemente, ya que entre otras cosas, para un píxel o característica dada en una imagen pueden existir multitud de candidatos en la otra imagen. Para resolver esta problemática, hemos de abordar las cuestiones:

- ¿Qué elementos deben considerarse para la correspondencia?
- ¿Qué restricciones reducen la ambigüedad del proceso?

Los elementos más comúnmente empleados en la literatura general son:

- Píxel: podemos usar las características del píxel: nivel de gris, niveles de gris de la vecindad (ventana), conjunto de descriptores invariantes.
- Borde: se usan las características de segmentos rectilíneos: longitud, orientación, nivel gris medio.
- Píxel + borde.

Para abordar el problema de la correspondencia se utilizan lo que se conocen como restricciones estereoscópicas, destacando las siguientes: *similitud*, *unicidad*, *continuidad de la disparidad*, *epipolaridad*, (Pajares y Cruz, 2007). La similitud se refiere al hecho de que dado un punto en la escena éste se proyectará con propiedades similares en la imagen, por ejemplo las hojas verdes de un árbol se proyectan en color verde en sendas imágenes. La unicidad indica que un punto en la escena se proyecta en un único punto en cada imagen (salvo oclusiones o fenómenos especulares). La continuidad de la disparidad indica que los píxeles de los objetos se situarán a distancias similares y en consecuencia a disparidad constante, por ejemplo los píxeles de un mismo tronco de árbol están todos ellos a la misma distancia del sistema estéreo y por tanto presentan la misma disparidad. Finalmente, la epipolaridad, se refiere a que las proyecciones de un mismo punto en la escena en sendas imágenes lo hacen siguiendo precisamente las denominadas líneas epipolares.

En el algoritmo de correspondencia que se presenta en este documento solo se han tenido en cuenta las restricciones de similitud y epipolaridad. La similitud no puede ser garantizada completamente en nuestro caso debido a los efectos de iluminación.

### ***Epipolaridad***

En Romero (2008) se describe el proceso calibración de cámaras estereoscópicas con lentes de “ojo de pez” y las ecuaciones que definen la línea epipolar. La figura 3.19 representa la proyección de un punto objeto del espacio en cada una de las dos imágenes captadas por la cámara. El punto viene descrito por las coordenadas  $(x_1, y_1, z_1)$  y  $(x_2, y_2, z_2)$  en los sistemas de coordenadas de las cámaras 1 y 2 respectivamente. Se supone que la orientación de las cámaras es tal que comparten el eje Y y el plano XY. Además,  $\alpha_1$  y  $\alpha_2$  son los ángulos de incidencia del haz del punto objeto respecto al eje óptico de cada una de las cámaras;  $(x_{1i}, y_{1i})$  y  $(x_{2i}, y_{2i})$  son las coordenadas de los puntos imagen para cada cámara; los puntos  $(x_1, y_1, z_1)$  y  $(x_{1i}, y_{1i})$  y el eje  $z_1$ , pertenecen al mismo plano, al igual que ocurre con los puntos  $(x_2, y_2, z_2)$  y  $(x_{2i}, y_{2i})$  y el eje  $z_2$ . De acuerdo con lo expresado previamente, podemos establecer las siguientes relaciones (Schwalbe, 2005):

$$\begin{aligned} \frac{x_{1i}}{y_{1i}} &= \frac{x_1}{y_1} \\ \frac{x_{2i}}{y_{2i}} &= \frac{x_2}{y_2} \end{aligned} \quad (3.1)$$

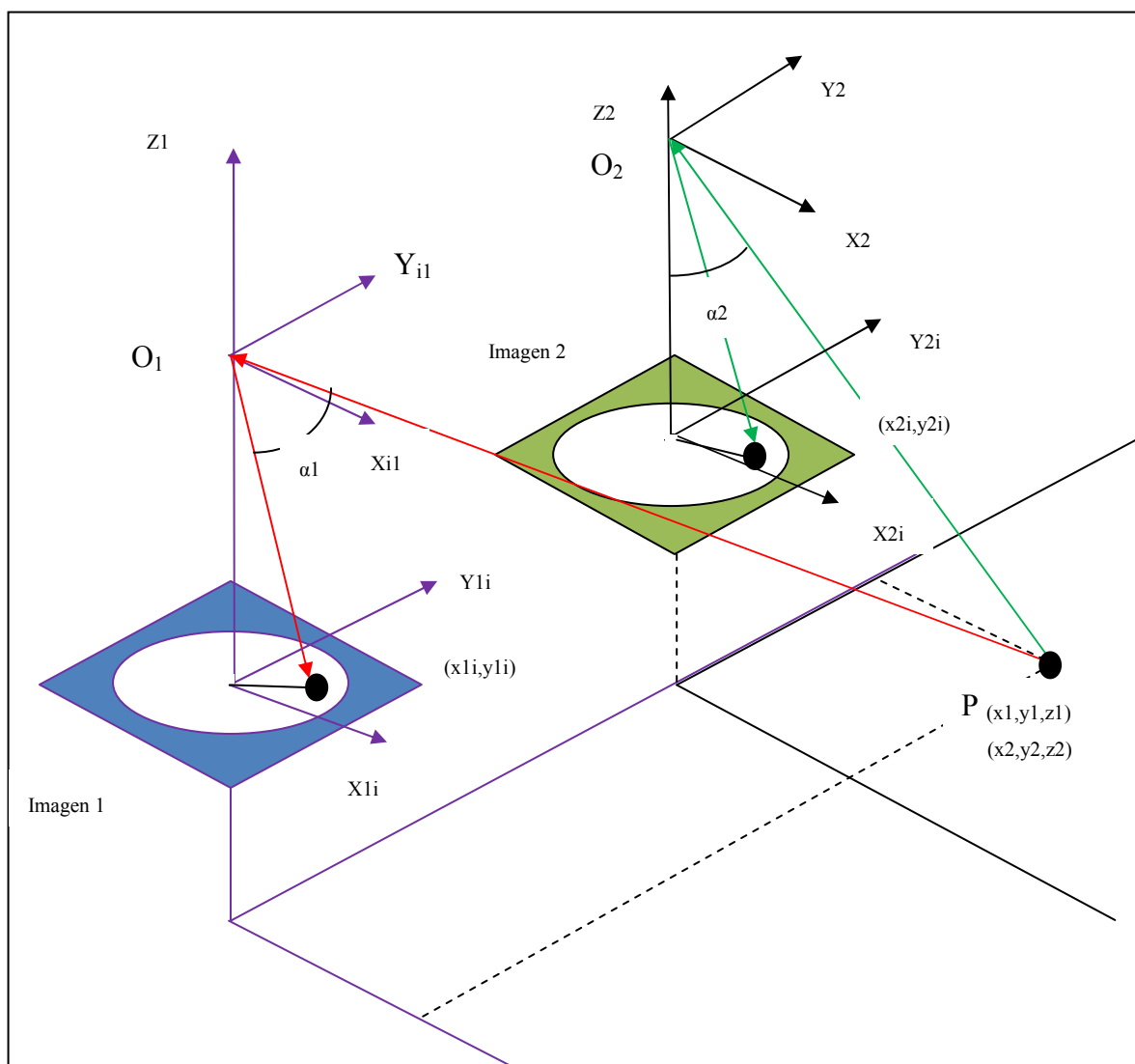


Figura 3.19 Relación geométrica de las coordenadas de un punto objeto y su proyección en el plano imagen de dos cámaras de un sistema estéreo

Dado un punto  $(x_{1i}, y_{1i})$  de la imagen 1, los puntos objeto del espacio cuya proyección coinciden en el mismo, determinan la recta  $\overline{O_1 P}$  en el espacio. Por otro lado, aplicando a cada uno de los puntos de la recta  $\overline{O_1 P}$  su proyección en la imagen 2, obtendremos una línea que deberemos recorrer para localizar el punto homólogo. Esta línea se llama epipolar y su determinación ayudará a resolver el problema de la correspondencia entre las dos imágenes restringiendo el problema a la consideración de un único grado de libertad.

La proyección de un punto objeto en la imagen dos viene dada por (Schwalbe, 2005):



$$\begin{aligned}
x_{2i} &= \frac{\frac{2.R}{\pi} . a \tan \left[ \frac{\sqrt{x_2^2 + y_2^2}}{z_2} \right]}{\sqrt{\left( \frac{y_2}{x_2} \right)^2 + 1}} \\
y_{2i} &= \frac{\frac{2.R}{\pi} . a \tan \left[ \frac{\sqrt{x_2^2 + y_2^2}}{z_2} \right]}{\sqrt{\left( \frac{x_2}{y_2} \right)^2 + 1}}
\end{aligned} \tag{3.2}$$

En la ecuación (3.2) se ha omitido el factor de distorsión de la lente, por simplicidad. Por otro lado, la transformación entre los sistemas de coordenadas de las cámaras 1 y 2 se puede expresar de forma simple mediante:

$$\begin{aligned}
x_2 &= x_1 \\
y_2 &= y_1 - y_{12} \\
z_2 &= z_1
\end{aligned} \tag{3.3}$$

donde  $y_{12}$  representa la separación entre las cámaras 1 y 2. Sustituyendo (3.3) en (3.2) y teniendo en cuenta la ecuación de la recta  $\overline{O_1P}$  definida por los posibles puntos objeto cuya proyección es  $(x_{1i}, y_{1i})$  tenemos,

$$\begin{aligned}
\frac{y_1}{x_1} &= \frac{y_{1i}}{x_{1i}} \\
\frac{z_1}{x_1} &= \frac{\sqrt{1 + \left( \frac{y_{1i}}{x_{1i}} \right)^2}}{\tan\left(\frac{\pi}{2R} \sqrt{x_{1i}^2 + y_{1i}^2}\right)}
\end{aligned} \tag{3.4}$$

A partir de lo anterior podremos calcular la línea trazada por los puntos de coordenadas  $(x_{2i}, y_{2i})$  obtenidos al recorrer la recta  $\overline{O_1P}$ . La línea así obtenida es la epipolar y servirá de guía para determinar la correspondencia de un punto de la imagen 1 en la imagen 2, tal y como se ha mencionado previamente.

### 3.3.1. Métodos basados en el área

Como ya se comentó previamente, los métodos basados en el área buscan la correspondencia entre las imágenes usando la correlación como medida de similitud. Teniendo en cuenta las restricciones anteriormente indicadas se ha probado la correspondencia a nivel de píxel y a nivel de ventana para aquellas regiones de la imagen pertenecientes a troncos.

Se han probado las medidas de similitud definidas en las ecuaciones (2.6), (2.7) y (2.8) sobre las regiones de la figura 3.20.

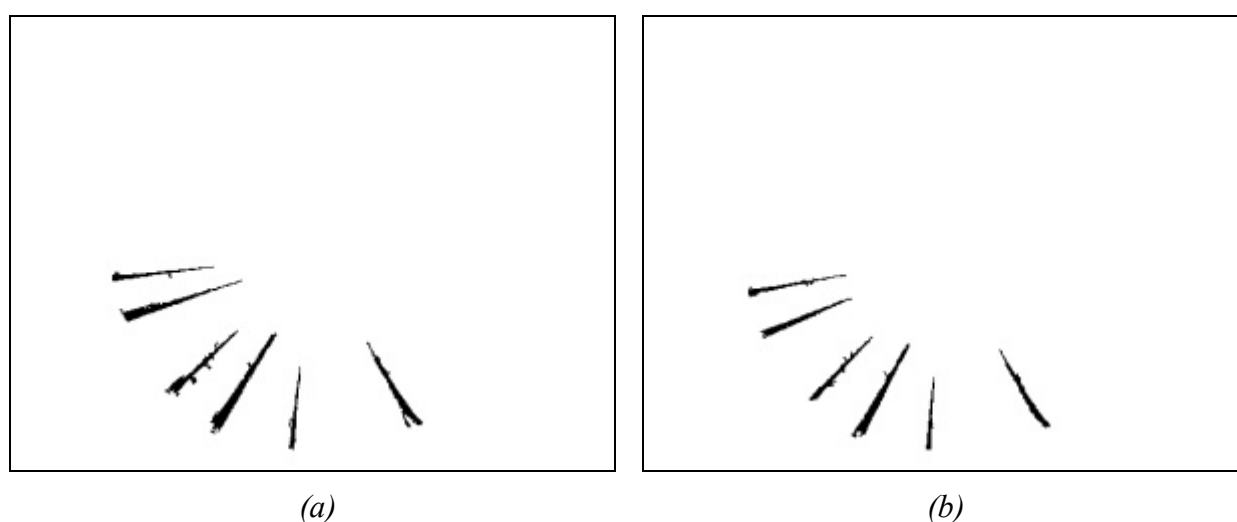


Figura 3.20 Regiones segmentadas a partir de un par de imágenes captadas con un sistema estéreo. Por cada píxel en la imagen 3.20(a) se busca su homólogo en la imagen 3.20(b) y obtenemos una diferencia de ángulo que se representa con un color. En la figura 3.21 se muestra los rangos de colores utilizados para representar las diferencias de ángulos.

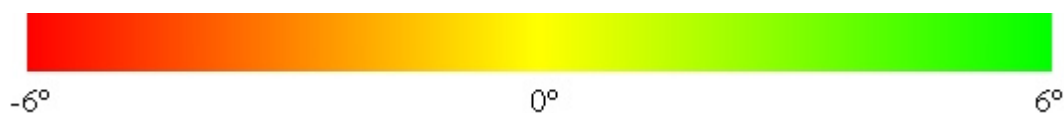


Figura 3.21 Rango de colores utilizados para representar la correspondencia.

#### ***Correlación cruzada***

En la figura 3.22 pueden verse los resultados obtenidos usando la correlación cruzada (2.6) y usando distintos tamaños de ventana sobre las imágenes de la figura 3.20.

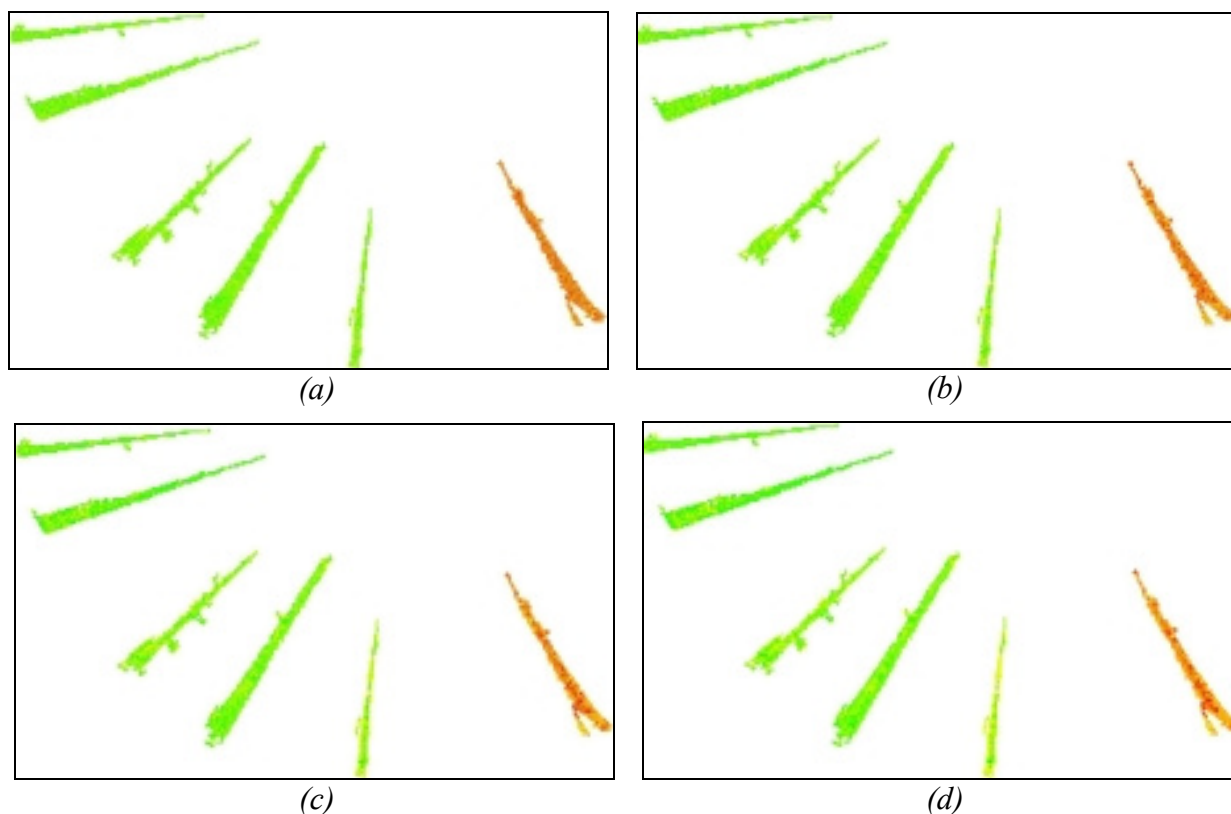


Figura 3.22 Correspondencia basada en el área usando correlación cruzada y distintos tamaños de ventana (N). (a) Imagen original. (b) Correspondencia con N=3. (c) Correspondencia con N=5. (d) Correspondencia con N=7. (e) Correspondencia con N=11.

Usando la correlación cruzada como función de similitud no tiene sentido usar un tamaño de ventana  $N=1$  ya que en ese caso siempre nos da como resultado 0.

Como puede observarse en la figura 3.21 las disparidades obtenidas para cada píxel no son continuas. Con esta función de similitud los resultados no son los esperados.

### ***Suma de diferencias absolutas***

En la figura 3.23 pueden verse los resultados obtenidos usando la suma de diferencias absolutas según la ecuación (2.7) y usando distintos tamaños de ventana sobre las imágenes de la figura 3.20.

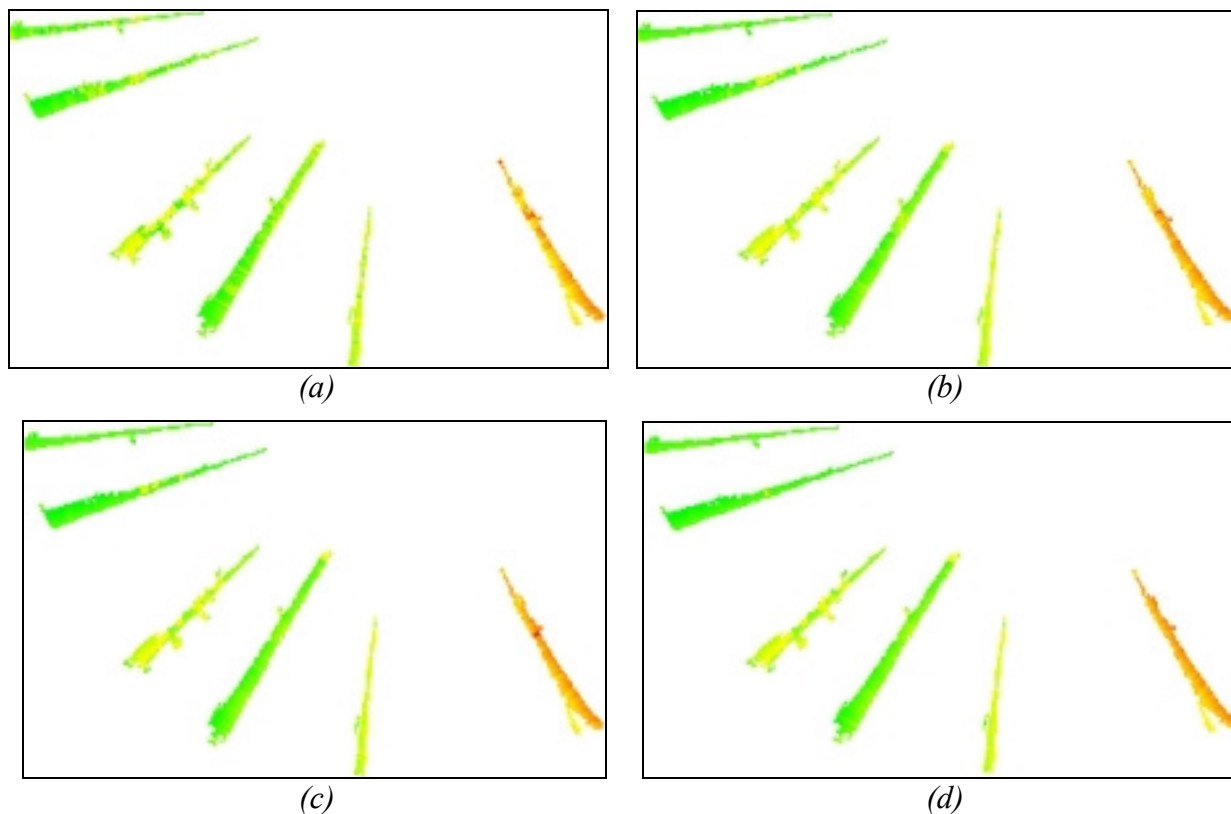


Figura 3.23 Correspondencia basada en el área usando suma de diferencias absolutas y distintos tamaños de ventana (N). (a) Imagen original. (b) Correspondencia con N=1. (c) Correspondencia con N=3. (d) Correspondencia con N=5. (e) Correspondencia con N=7.

Cuando se usa un tamaño de ventana  $N=1$ , en realidad se están haciendo comparaciones a nivel de píxel.

Como puede observarse en la figura 3.23 las disparidades obtenidas tienen una mayor continuidad que con correlación cruzada salvo en algún tronco particular. Podemos ver que en los árboles situados cerca de  $270^\circ$  la diferencia de ángulos es cercana a  $0$  mientras que en ángulos cercanos a  $180^\circ$  la diferencia es cercana a  $6^\circ$ .

### ***Suma de los cuadrados de las diferencias***

Por último, en la figura 3.24 pueden verse los resultados obtenidos usando la suma de los cuadrados de las diferencias según la ecuación (2.8) como medida de similitud y con distintos tamaños de ventana.

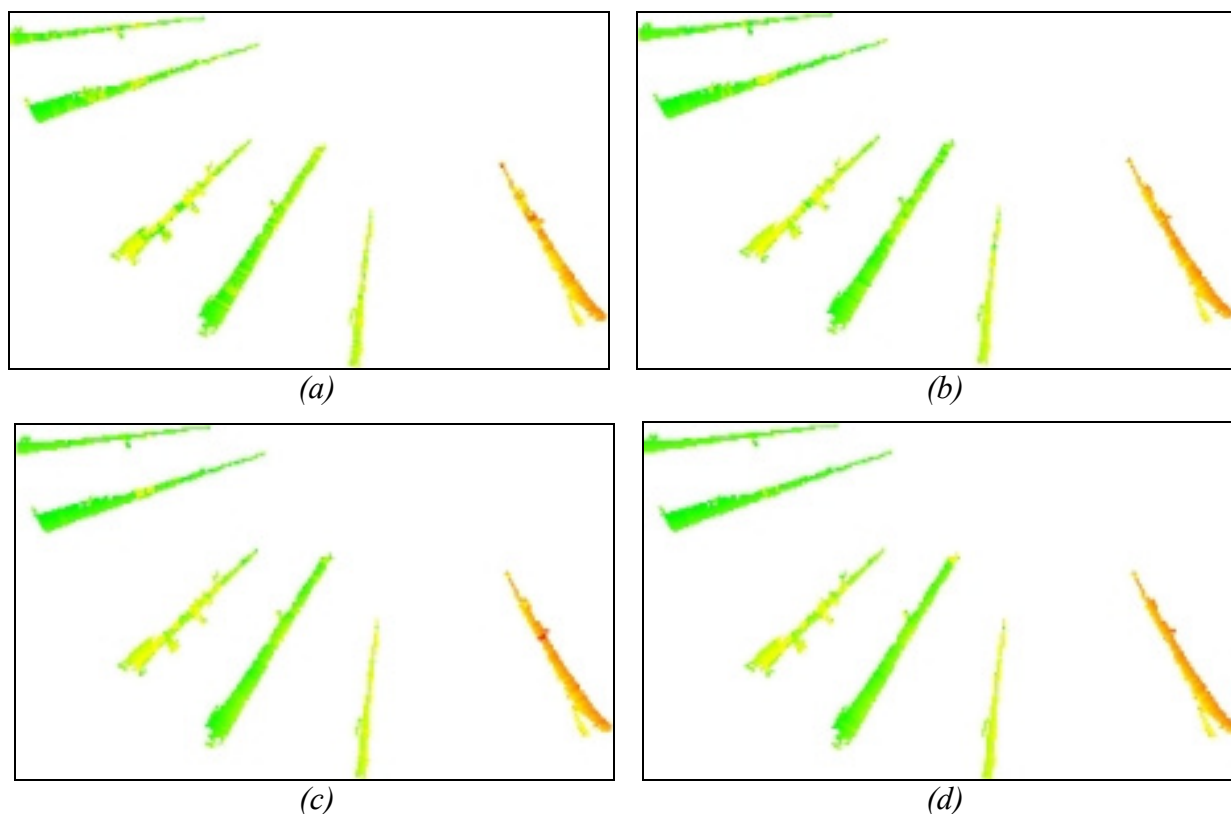


Figura 3.24 Correspondencia basada en el área usando suma de cuadrados de las diferencias y distintos tamaños de ventana (N). (a) Imagen original. (b) Correspondencia con  $N=1$ . (c) Correspondencia con  $N=3$ . (d) Correspondencia con  $N=5$ . (e) Correspondencia con  $N=7$ .

Los resultados obtenidos con esta función de similitud proporcionan resultados similares a los obtenidos con la suma de diferencias absolutas.

Como puede observarse con tamaños de ventana muy pequeños no se obtienen resultados satisfactorios debido a que no se captura suficiente información de la textura.

#### *Algoritmo de correspondencia basado en la comparación de rectas*

Como ya se ha comentado previamente, la imagen derecha capturada con el sistema estéreo presenta cierta desviación con respecto a la imagen izquierda. En la figura 3.25 se muestran dos imágenes capturadas por el sistema estéreo sobre las que se han dibujado ciertos ángulos permitiendo ver con mayor claridad este desfase de ángulos.

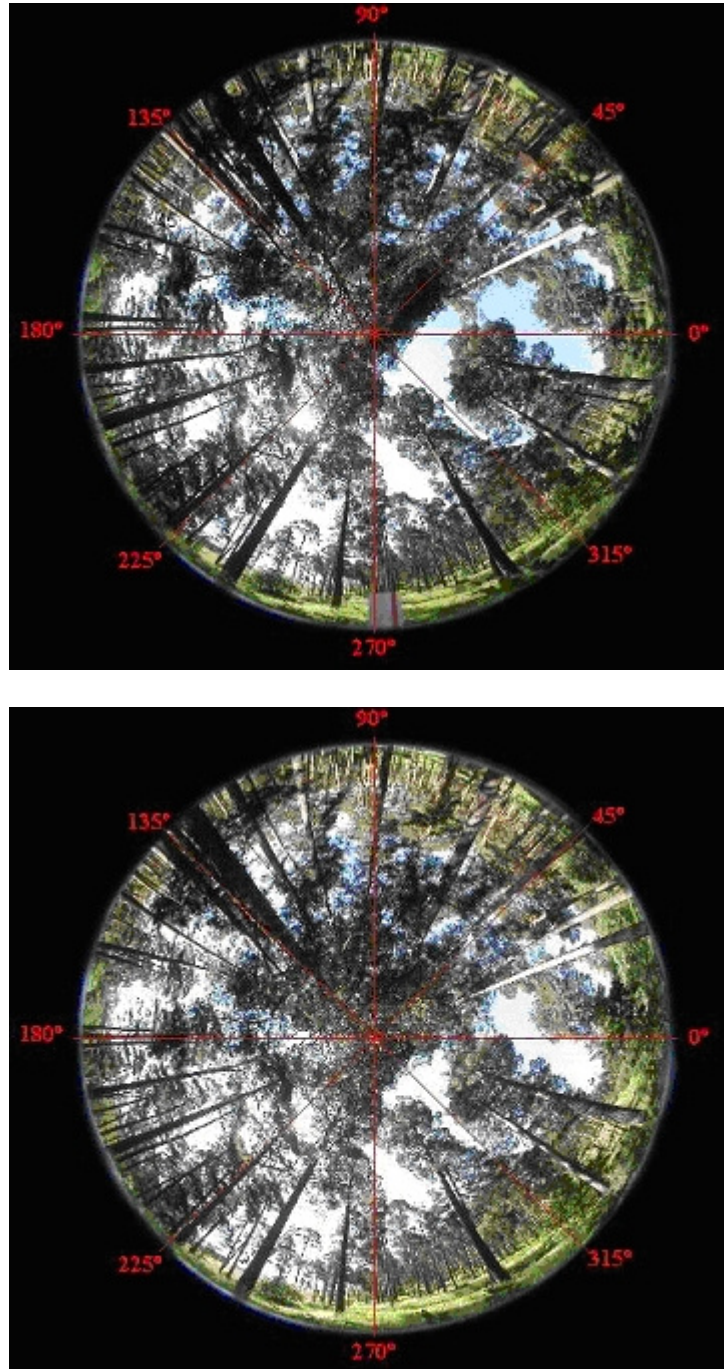


Figura 3.25 Par estéreo con desfases.

A la vista de estas imágenes se observa que, el hemisferio derecho de la imagen derecha tiene un desfase negativo mientras que el hemisferio izquierdo tiene un desfase positivo respecto a la imagen izquierda.

Aprovechando este desfase se ha diseñado un algoritmo que, a partir de una recta con un ángulo  $\alpha$  en la imagen izquierda, busca la recta homóloga con un ángulo  $\beta$  en la imagen derecha. A la hora de comparar las rectas realmente estaremos comparando puntos sobre



dicha recta por lo que será necesario decidir cuántos puntos seleccionaremos y cómo distribuirlos a lo largo de la recta. En la figura 3.26 podemos ver la búsqueda de una recta en la imagen izquierda (a) sobre la imagen derecha (b).



(a)



(b)

Figura 3.26 Correspondencia usando rectas.

Tendremos que definir entonces el valor del desfase máximo que usaremos a la hora de buscar la correspondencia y en qué sentido de giro aplicar dicho desfase. Como ya hemos comentado, el hemisferio derecho de la imagen derecha tiene un desfase negativo y el

hemisferio izquierdo tiene un desfase positivo, por lo que cuando estemos buscando la correspondencia en ángulos mayores que  $0^\circ$  y menores que  $90^\circ$  o en ángulos mayores que  $270^\circ$  y menores de  $360^\circ$  aplicaremos un desfase negativo. En ángulos mayores que  $90^\circ$  y menores que  $270^\circ$  aplicaremos un desfase positivo. En los ángulos  $90^\circ$  y  $270^\circ$  se presupone que no hay desfase.

En la figura 3.26 podemos ver las rectas usadas en la búsqueda teniendo en cuenta un desfase máximo de  $6^\circ$ . En la imagen izquierda se está usando un radio con un ángulo de  $239^\circ$  buscándose en la imagen de la derecha en el rango  $239^\circ$ - $245^\circ$ . En este caso el algoritmo de correspondencia nos dice que la recta más similar a la de la imagen izquierda es la que tiene un desfase de  $4^\circ$ , es decir, la quinta recta. Este método nos permite establecer la correspondencia entre árboles obteniendo resultados aceptables pese a las zonas con hojas que existen en las imágenes. Este proceso se repite para los  $360^\circ$  de la circunferencia.

Hasta este punto es el método que ha sido probado sin demasiado éxito. Como posible mejora se propone, en el futuro, realizar un suavizado de los desfases obtenidos para corregir posibles errores en la correspondencia.

Las decisiones más importantes que debemos entonces tomar son:

- el desfase máximo entre las dos imágenes,
- cuántos puntos y cómo distribuirlos sobre las rectas que se van a comparar.

Para conocer el desfase máximo tendremos que saber la distancia mínima a la que se encuentran los troncos que queremos reconocer y a partir de ese dato podemos saber el desfase máximo que debemos recorrer.

Para el segundo aspecto, podemos elegir puntos equidistantes o puntos aleatorios. Ambas opciones tienen ventajas e inconvenientes.

El problema que presenta este algoritmo de correspondencia es que en regiones con texturas muy similares (zonas de hojas, zonas de sólo cielo,...) no se asegura la correcta correspondencia. En la figura 3.27 se ilustra este caso.



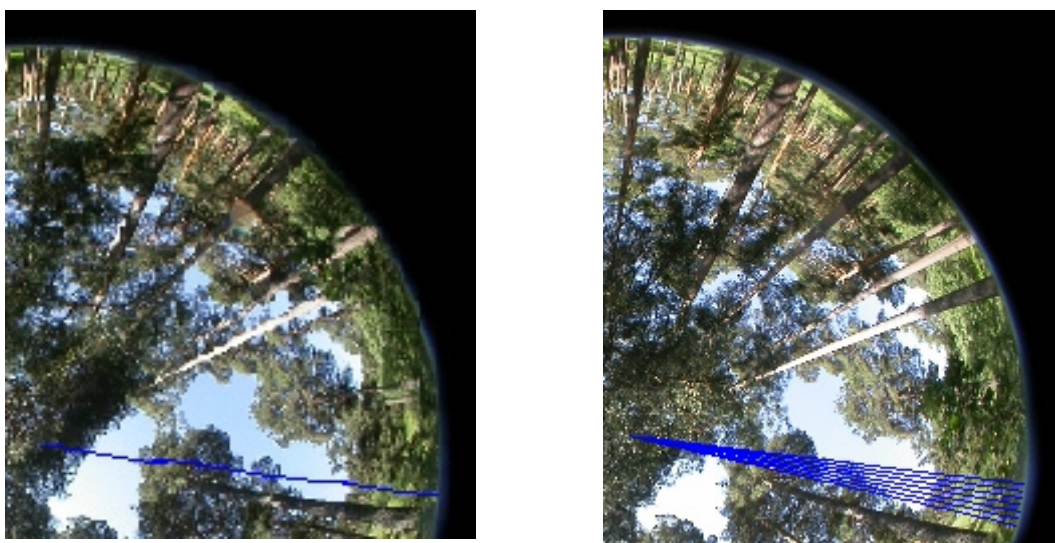


Figura 3.27 Correspondencia usando recta.

### 3.3.2. Métodos basados en características

Los métodos basados en características buscan la correspondencia ayudándose de características específicas de las propias imágenes. En el caso de las imágenes forestales la complejidad que presentan es un inconveniente para la obtención de características discriminantes. Es por este motivo, por el que se ha decidido extraer características a partir de las imágenes ya segmentadas.

Entre las características que pueden usarse para buscar la correspondencia están las esquinas, puntos de borde y segmentos de línea como ya se comentó en capítulo dos. El número de esquinas que obtendríamos en nuestras imágenes sería demasiado alto debido a la irregularidad de los troncos. Por lo tanto, las características más prometedoras podrían ser puntos de borde y segmentos de líneas.

Si aproximamos los troncos de los árboles a segmentos de línea, podremos obtener su longitud, su orientación, las coordenadas del punto medio y la media de intensidad en la línea como características para comparar en el algoritmo. En la figura 3.28 se ilustra la región segmentada perteneciente a un tronco (a) y el segmento que lo aproxima.

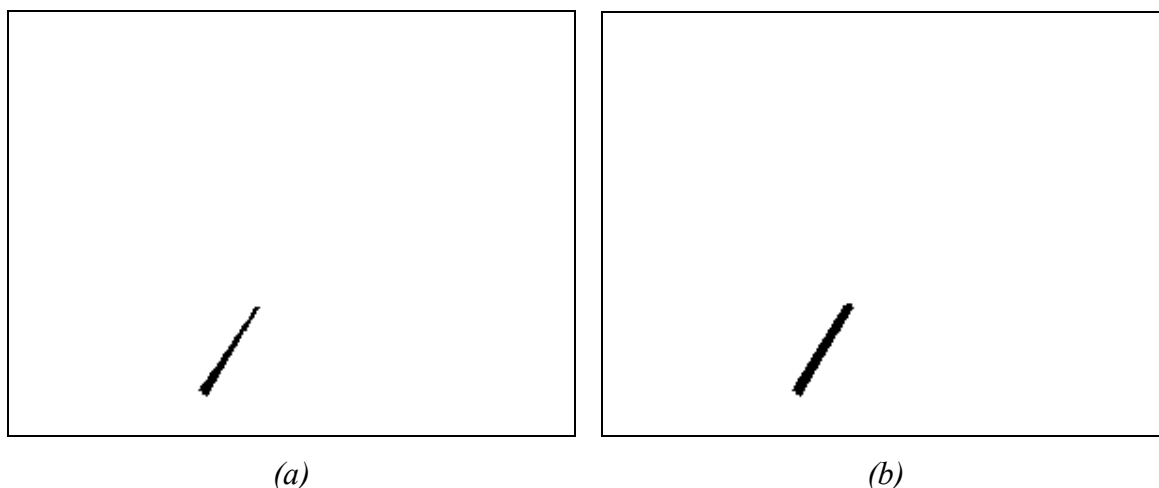


Figura 3.28 (a) Tronco segmentado en Imagen izquierda. (b) Segmento de línea que aproxima al tronco segmentado.

A partir del segmento de línea de la figura 3.28 (b) podemos obtener sus características y buscar el segmento mas semejante en la imagen derecha captada por el sistema estéreo.

Para obtener puntos de borde podemos usar diversos operadores de primera y segunda derivada. La teoría matemática del gradiente, basado en el concepto de primera derivada, se aplica satisfactoriamente a la extracción de bordes. A partir de este concepto se obtienen una serie de operadores de borde basados en la primera derivada, a saber Sobel, Prewitt, Roberts, Kirsch, Robinson, o Frei-Chen. Los fundamentos de los mismos se pueden encontrar en Gonzalez y Woods (1993), Trucco y Verri (1998), Haralick y Sapiro (1992), Fu y col (1988), Umbaugh (1998), Pratt (1991), Pitas (2000). Una aplicación concreta de los métodos basados en el gradiente se debe a Canny (1986, 1987), que deriva un algoritmo para la extracción de bordes, adelgazamiento y cierre de los mismos, el cual puede encontrarse con detalle en Trucco y Verri (1998).

Respecto de los operadores segunda derivada destacan la Laplaciana cuyos fundamentos los encontramos en González y Woods (1993), Fu y col. (1988); la Laplaciana de la Gaussiana en Marr (1982, 1985), Huertas y Medioni (1986), Grimson y Hildreth (1985), Ramaswamy y col (1990), Trucco y Verri (1998), o Jähne (1995) y la diferencia de Gaussianas en Jähne (1995).

No se han llegado a probar la correspondencia basada en características por lo que no se profundizará en este tipo de correspondencia y se deja como objetivo para trabajos futuros.

### **3.3.3. Resumen de los métodos de correspondencia**

A la vista de los resultados mostrados previamente, se ha comprobado que ventanas con tamaños grandes proporcionan mejores resultados a costa de mayores tiempos de proceso. Respecto a las funciones de similitud, las que han presentado un mejor funcionamiento han sido la suma de diferencias absolutas (2.7) y la suma de los cuadrados de diferencias (2.8). Ha sido probado un algoritmo que a partir de una recta con un ángulo  $\alpha$  en la imagen izquierda, busca la recta homóloga con un ángulo  $\beta$  en la imagen derecha. No se han llegado a resultados satisfactorios aunque se establece como base de referencia para futuros trabajos.

La correspondencia basada en características no se ha implementado por el amplio trabajo que por sí mismo implica, dejándose para futuros trabajos.



# Capítulo 4

## 4. Análisis de resultados

### 4.1. Objetivos del análisis

En este capítulo se hará verificación y validación de las estrategias propuestas. Puesto que, como ya se ha mencionado, no existe mucha investigación específica relevante en este área, no resulta posible comparar los métodos propuestos con otros ya existentes, por la inexistencia de los mismos. Por otra parte y dado el grado de avance del proyecto, tampoco es posible obtener resultados cuantitativos verificables por los técnicos forestales del INIA y la ETSIM debido a que falta completar el proceso mediante el análisis estereoscópico, que se plantea para futuros trabajos de investigación. Por tanto, el principal objetivo será analizar las ventajas e inconvenientes de cada uno de los métodos propuestos.

En cuanto a la precisión de los resultados, es difícil establecer cuál es el mejor en determinadas ocasiones, ya que el único método conocido de validación es el ojo humano que es una medida subjetiva.

#### *Descripción de las imágenes utilizadas*

Se dispone de un conjunto de imágenes adquiridas por los técnicos forestales de la ETSIM. Se trata de imágenes digitales estéreo comprimidas en formato JPG y con una dimensión de 2272x1704 píxeles. Los sistemas estéreo producen un par de imágenes con un pequeño desplazamiento entre una y otra, este sistema es similar al que usamos los seres humanos en el que captamos dos imágenes con un pequeño desplazamiento para poder calcular las distancias en la escena 3D.

Seleccionaremos cuatro imágenes del conjunto inicial para realizar las pruebas con los distintos algoritmos de extracción de troncos. A este conjunto lo llamaremos conjunto de pruebas.

A continuación, se muestran en las figuras 4.1 y 4.2 un par de imágenes del conjunto de pruebas con sus correspondientes pares estéreo. Tal y como se ha descrito en capítulos previos, estas imágenes entrañan una dificultad evidente por la gran cantidad de elementos presentes en las mismas y que de manera muy frecuente se cruzan unas con otros. Atendiendo al problema, podríamos clasificar las texturas existentes como:

- Troncos de los árboles, habitualmente en marrón
- Hojas de los árboles, normalmente en verde
- Cielo, representado por un color que va desde el negro al azul claro
- Suelo, puede ser de cualquier color, pero normalmente será marrón o verde



(a)



(b)

Figura 4.1 (a) Imagen izquierda; (b) Imagen derecha



(a)



(b)

Figura 4.2 (a) Imagen izquierda; (b) Imagen derecha

## 4.2. Análisis de resultados

### 4.2.1. Criterio de validación

Debido a la mencionada inexistencia de métodos conocidos para la comparación se hace necesario establecer para cada uno de los algoritmos un criterio de comprobación. El objetivo por tanto, se centra ahora en establecer un procedimiento para poder llevar a cabo el proceso de validación.

Queda también, por lo mencionado previamente, totalmente fuera del alcance de este trabajo el estudio o creación de un test que obtenga una medida cualitativa de cada uno de los algoritmos. De esta manera, trataremos de centrar nuestro análisis en las diferencias existentes en cada uno de los algoritmos, explicando sus ventajas y sus inconvenientes. La calidad de los resultados obtenidos se evalúa de forma cualitativa atendiendo a criterios subjetivos de los técnicos forestales.

A continuación, analizaremos los métodos de segmentación y correspondencia usados describiendo sus principales ventajas e inconvenientes.

#### *Segmentación*

Con el proceso de segmentación se quiere simplificar la imagen para luego generar una mejor correspondencia. Se han probado los siguientes métodos de segmentación:

- Segmentación por umbralización. Esta técnica es muy limitada. Sólo será válida cuando el color de las hojas, cielo, suelo y troncos sean distintos y se puedan tener umbrales capaces de separarlos. Esta técnica puede ser utilizada conjuntamente con otras técnicas para realizar un preprocesamiento, eliminando así ciertos elementos innecesarios del resto de la imagen.
- Segmentación basada en regiones:
  - Crecimiento de regiones. Como se describió en la sección 3.2.2.1, se puede aplicar un método supervisado o no supervisado para la selección de semillas. El primer método se comporta mejor que el segundo salvo en aquellos casos en que un tronco esté formado por regiones muy distintas. El

inconveniente es que el método deja de ser completamente automático, requiriendo la presencia de un experto que seleccione las semillas.

- División y Fusión. Este método funciona bien cuando las regiones de troncos sean uniformes y el resto de regiones no o presenten intensidades muy elevadas como ocurre en el caso de las texturas del tipo cielo.
- Watershed. Esta técnica es buena para imágenes simples. Cuando la imagen es compleja se produce sobre-segmentación sin obtener unos resultados esperanzadores.

### ***Correspondencia***

Se han probado técnicas de correspondencia basadas en el área y en las características. Si bien, sólo se ha probado el primer tipo de estrategias, dejando para futuros trabajos la correspondencia basada en características. Respecto a la correspondencia basada en el área, se han utilizado distintos tamaños de ventana y funciones de similitud. Se ha comprobado que ventanas con tamaños grandes funcionan mejor que con tamaños pequeños, eso sí incrementando el coste computacional. Respecto a las funciones de similitud, las que han presentado un mejor funcionamiento han sido la suma de diferencias absolutas, ecuación (2.7) y la suma de los cuadrados de diferencias ecuación (2.8). También ha sido probado un algoritmo que a partir de una recta con un ángulo  $\alpha$  en la imagen izquierda, busca la recta homóloga con un ángulo  $\beta$  en la imagen derecha. No se han llegado a resultados satisfactorios aunque se establecen las bases para futuras mejoras e investigaciones relacionadas.



## 5. Conclusiones y trabajo futuro

### 5.1. Sobre resultados

Se han propuesto en este trabajo varias aproximaciones para tratar de resolver el problema de la identificación de los troncos de los árboles en imágenes boscosas.

Se han propuesto distintos métodos ampliamente detallados en la literatura pero no aplicados previamente en el campo de las imágenes forestales. Algunos de estos métodos han resultado útiles para casos concretos y otros se han descartado. Con los resultados obtenidos por los distintos métodos probados se ha llegado a la conclusión de que ninguno de los métodos ofrece una solución válida para todas las posibles imágenes capturadas.

El trabajo desarrollado ha sido útil para familiarizarse con el desarrollo de trabajos de investigación y en un campo concreto como el de la visión artificial.

Los resultados obtenidos hacen que este trabajo sea una buena base de partida para posteriores experimentos en este campo.

Se encontraron varios problemas durante los estudios de los distintos métodos aplicados a la resolución del problema. Sin embargo, los principales fueron la extrema dificultad de las imágenes y la compresión de las mismas en el formato JPEG con pérdidas, lo que supone una pérdida de información muy valiosa. La compresión en las imágenes hace que se pierda gran parte de la información del color de las imágenes digitales. Esta compresión produce pérdida de información en algunas texturas que nos sería útil para distinguir, por ejemplo, los troncos de las hojas. Tal vez, disponiendo de esta información y en conjunción con alguno de los métodos propuestos se puedan optimizar los resultados.

Observando las imágenes capturadas puede comprenderse la dificultad que presenta automatizar la tarea de separar los troncos del resto de la imagen siendo incluso complicado en ocasiones para el ojo humano.

## 5.2. Cumplimiento de los objetivos propuestos

En relación a los objetivos generales expresados en la sección 1.2, cabe añadir que se han cumplido las expectativas previstas.

Respecto de los específicos, su grado de cumplimiento se concreta como sigue:

1. Se han estudiado una gran variedad de métodos en prácticamente la mayoría de las áreas relevantes. Esto ha permitido adquirir un conocimiento global bastante amplio del campo.
2. Se han utilizado imágenes estéreo de zonas boscosas
3. Se han propuesto e implementado varios métodos que dan aproximaciones más o menos precisas a la resolución del problema
4. Se proponen líneas de investigación futuras

## 5.3. Trabajos futuros

Como trabajos futuros surgidos durante la investigación llevada a cabo se han identificado los siguientes:

- Análisis de otras representaciones del color distintas al formato RGB. En Gonzalez y Woods (1993) encontramos un estudio exhaustivo y tratamiento del color.
- La precisión del sistema podría mejorar
- Desarrollar un método de correspondencia basado en características para no prestar atención únicamente a la información del color. Este método no ha llegado a implementarse dada la amplitud de ese campo.
- Mejora de los factores de iluminación. Hay casos en que los troncos presentan el mismo color que el cielo debido a cuestiones de iluminación. Otras imágenes han sido captadas con escasa iluminación por lo que el rango de colores ha quedado reducido a un rango de valores de intensidad muy pequeño.

## 6. Bibliografía

1. Anuta, P. F. (1969). Digital registration of multispectral video imagery, *Soc. Photo-Opt. Instrum. Eng. J.*, 7, 168-175.
2. Ayache, N. and Faverjon, B. (1987). Efficient Registration of stereo images by matching graph descriptions of edge segments, *Int. J. Computer Vision*, 2, 107-131.
3. Barnea, D.I. and Silverman, H.F. (1972). A Class of Algorithms for Fast Digital image Registration. *IEEE Trans. Computers*, 21, 179-186.
4. Beucher, S., and Lantuéjoul, C. (1979). Use of watersheds in contour detection. In *Proc. International Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation, Rennes, September*.
5. Boyer, K.L. and Kak, A.C. (1988). Structural stereopsis for 3-D vision. *IEEE Trans. Pattern Anal. Machine Intell.*, 10, 144-146.
6. Canny, J. (1986). A computational Approach to Edge detection, *IEEE Trans. Patt. Anal. Machine Intell.* 8, 679-698.
7. Canny, J. (1987). A computational Approach to Edge Detection, In Fischler, M.A. and Firschein, O. (Eds.). *Readings in Computer Vision*, Morgan Kaufmann, Los Altos: CA, pp. 184-203.
8. Coiras, R.T. Hanson, A.R. Riseman, M.R. and Schlitz, H. (1995). Automatic extraction of buildings and terrain from aerial images. In A. Grün, O. Kübler and P. Agouris (Eds.) *Automatic Extraction of Man Made Objects Aerial and Space Images*, Birkhäuser-Verlag, Basel, pp. 169-178.

9. Cruz, J.M., Pajares, G. and Aranda, J. (1995a). A Neural Network approach to the stereovision correspondence problem by unsupervised learning. *Neural Networks*, vol. 8, no. 5, pp. 805-813.
10. Cruz, J.M., Pajares, G., Aranda, J and Vindel, J.L.F. (1995b). Stereo Matching technique based on the perceptron criterion function. *Pattern Recognition Letters*, vol. 16, pp. 933-944.
11. Davies, E.R. (1997). *Machine Vision: Theory, Algorithms, Practicalities*. Academic Press, London.
12. Digabel, H., and Lantuéjoul, C. Iterative algorithms. In *Actes du Second Symposium Européen d'Analyse Quantitative des Microstructures en Sciences des Matériaux, Biologie et Médecine, Caen, 4-7 October 1977* (1978), J.-L. Chermant, Ed., Riederer Verlag, Stuttgart, pp. 85-99.
13. Fu, K.S., González, R.C. y Lee, C.S.G. (1988). *Robótica: Control, detección, Visión e Inteligencia*. McGraw-Hill, Madrid.
14. Gonzalez, R.C. and Woods, R.E. (1993). *Digital Image Processing*, Addison-Wesley, Reading, MA.
15. Grimson, W.E.L. (1985). Computational experiments with a feature-based stereo algorithms, *IEEE Trans. Pattern Anal. Mach. Intel.* 7, 17-34.
16. Grimson, W.E.L. and Hildreth, E.C.. (1985). Comment on digital step edges from zero-crossings of second directional derivatives Computational *experiments with a feature-based stereo algorithm*, *IEEE Trans. Pattern Anal. Mach. Intell.* 7, 121-126.
17. Haralick, R.M. and Shapiro, L.G. (1992). *Computer and Robot Vision*, vol. I, Addison Wesley, Reading, Ma.

18. Huertas, A. and Medioni, G. (1986). Detection of Intensity Changes with Subpixel Accuracy Using Laplacian-Gaussian Masks. *IEEE Trans. Pattern Anal. Mach. Intell.* 8(5), 651-664.
19. Kim, Y.C. and Aggarwal, J.K. (1987). Positioning 3-D Objects using stereo images. *IEEE J. Robotics and Automation*, 3(4), 361-373.
20. Kim, D.H., Choi, W.Y. and Park, W.Y. (1992). Stereo matching technique based on the theory of possibility. *Pattern Recognition Letters*, 13, 735-744.
21. Khotanzad, A., Bokil, A. and Y.W. Lee, Y.W. (1993). Stereopsis by Constraint Learning Feed-Forward Neural Networks, *IEEE Trans. Neural Networks*. 4, 332-342.
22. Krotkov, E., Henriksen, K. and Kories, R (1990). Stereo ranging with verging cameras. *IEEE Trans. Pattern Anal. Machine Intell.*, 12(12), 1200-1205.
23. Lee, S.H. and Leou, J.J. (1994). A dynamic programming approach to line segment matching in stereovision. *Pattern Recognition*, 27, 961-986.
24. Lew, M.S, Huang, T.S. and Wong, K. (1994). Learning and Feature Selection in Stereo Matching. *IEEE Tran. Pattern Anal. Machine Intell.*, 16(9), 869-881.
25. Li, Z.N. (1994). Stereo Correspondence Based on Line Matching in Hough space using Dynamic programming. *IEEE Trans. Sys. Man Cybern*, 24(1), 144-152.
26. Lutsiv, V.R. and Novikova, T.A. (1992). On the Use of a Neurocomputer for Stereoimage Processing. *Pattern Recognition and Image Analysis*. 2, 441-444.
27. Maravall, D. and Fernández, E. (1992). Contribution to the matching problem in stereovision, *Proc. 11th IAPR: International Conference in Pattern Recognition, The Hague*, pp. 441-414
28. Maravall, D. (1993). *Reconocimiento de Formas y Visión Artificial*. RA-MA.

29. Marr, D. and Poggio, T. (1976). Cooperative computation of stereo disparity, *Science*. 194, 283-287.
30. Marr, D. and Poggio, T. (1979). A computational theory of human stereovision. *Proceedings Royal Society of London*, B 207,301-328.
31. Marr, D. (1982). *Vision*, W.H. Freeman & Co., San Francisco.
32. Marr, D. (1985). *La Visión*, Alianza Editorial, Madrid.
33. Marshall, A.D. and R.R. Martin (1993). *Computer Vision models and Inspection*, World Scientific Series in Robotics and Automated Systems, vol. 4, World Scientific, London,.
34. Medioni, G. and Nevatia, R. (1984). Matching Images Using Linear Features, *IEEE Trans. Pattern Anal. Machine Intell.* 6(6), 675-685.
35. Medioni, G. and Nevatia, R. (1985). Segment Based Stereo Matching, *Computer Vision, Graphics, and Image Processing*, 31, 2-18.
36. Mousavi, M.S. and Schalkoff, R.J. (1994). ANN Implementation of Stereo Vision using a multi-Layer Feedback Architecture, *IEEE Trans. On Systems, Man, and Cybernnetics*, 24(8), 1220-1238.
37. Nasrabadi, N.M. (1992). A Stereo vision technique using curve-segments and relaxation matching. *IEEE Trans. Patt. Anal. Machine Intell*, 14(5), 566-572.
38. Newman, T.S. and Jain, A.K. (1995). *A survey of automated visual inspection*. *Computer Vision and Image Understanding*. 61(2), 231-262.
39. Ohta, Y. and Kanade, T. (1985). Stereo by intra- and inter-scanline search. *IEEE Trans. Pattern Anal. Machine Intell.*, 7(2), 139-154.
40. Otsu, N. (1979). A threshold selection method from gray-level-histograms. *IEEE Trans. System Man and Cybernetics*. SMC-9, 62-66.

41. Pajares, G., Pereira, R., Rives, J. y Díez, J.A. (1993). *Correspondencia Difusa en Visión Estereoscópica*. En III Congreso Español sobre tecnologías y Lógica Fuzzy (S. Barro y A. Sobrino, Eds.), pp. 303-310, Dpto. Electrónica de la Computación, Santiago de Compostela.
42. Pajares, G. and de la Cruz, J.M. (1997a). Synthesis of Learning Methods Applied to Stereovision Matching. *Proc. 5th European Congress on Intelligent Techniques and Soft Computing in Aachen, Germany*, pp. 1970-1974.
43. Pajares, G. and de la Cruz, J.M. (1997b). Fuzzy Criteria Applied to Stereovision Matching. *Proc. 5th European Congress on Intelligent Techniques and Soft Computing in Aachen, Germany*, pp. 1888-1892.
44. Pajares, G. and de la Cruz, J.M. (1997c). Synthesis of Learning methods in Stereovision matching. *VII Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes*. Barcelona.
45. Pajares, G., de la Cruz, J.M. and Aranda, J. (1998a). Relaxation by Hopfield Network in Stereo Image Matching. *Pattern Recognition*, 31(5), 561-574.
46. Pajares, G., de la Cruz, J.M. and Aranda, J. (1998b). Stereo Matching based on the Self-Organizing Feature-Mapping Algorithm. *Pattern Recognition Letters*, 19, 319-330.
47. Pajares, G., de la Cruz, J.M. and López-Orozco, J.A. (1998c). *Improvising Stereovision Matching through Supervised Learning*, Pattern Analysis and Applications, 1, 105-120.
48. Pajares, G., de la Cruz, J.M. and López-Orozco, J.A. (1998d). Pattern Recognition Learning Applied to Stereovision Matching. In A. Amin, D. Dori, P. Pudil and H. Freeman (Eds.), *Advances in Pattern Recognition*, Springer-Verlang, Berlin, pp. 997-1004.

49. Pajares, G., de la Cruz, J.M. and López-Orozco, J.A. (1999). Stereo Matching Using Hebbian Learning. *IEEE Transactions on Systems, Man and Cybernetics*, 29(4) Part B, 553-559.
50. Pajares, G. and de la Cruz, J.M. (2007). *Visión por Computador: Imágenes Digitales y Aplicaciones*, RA-MA.
51. Pitas, I. (2000). *Digital Image Processing Algorithms and Applications*. Wiley, New York.
52. Pratt, W.K. (1991). *Digital Imagen Processing*, John Willey and Sons, New York.
53. Ramaswamy, R., Morris. D.T. and Aspinall, D. (1990). Detection edges in an image using powers-of-two coefficients. *IEE Proc. Vol. 137, Pt. E, n° 3*.
54. Romero, M. A. (2008). *Calibración de cámaras estereoscópicas con lentes de “ojo de pez” para análisis de texturas forestales*. Programa de doctorado Ingeniería de Sistemas y Automática, Proyecto del periodo de investigación.
55. Rubio, P. (1993). RP: un algoritmo eficiente para la búsqueda de correspondencias en visión estereoscópica. *Informática y Automática*, 26, 5-15.
56. Ruycheck, Y. and Postaire, J.G. (1996). A neural network algorithm for 3-D reconstruction from stereo pairs of linear images, *Pattern Recognition Letters*, 17, 387-398.
57. Scharstein, D. and Szeliski, R. (2001). *Intl Journal of Computer Vision*, 47(1), 7-42.
58. Schwalbe, E., (2005). Geometric Modelling and Calibration of Fisheye Lens Camera Systems. Proc. 2<sup>nd</sup> Panoramic Photogrammetry Workshop, Int. Archives of Photogrammetry and Remote Sensing, Vol. 36, Part 5/W8,
59. Shirai, Y. (1987). *Three-dimensional Computer Vision*. Springer-Verlag, Berlín.



60. Sonka, M., Hlavac, V. and Boyle, R. (1995). *Image Processing, Analysis and Machine Vision*, Chapman-Hall, Cambridge.
61. Tanaka, S. and Kak, A.C. (1990). A Rule-Based Approach to Binocular Stereopsis. In R.C. Jain and A.K. Jain (Eds.) *Analysis and Interpretation of Range Images*. Springer-Verlang, New York.
62. Trucco, E. and Verri, A. (1998). *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall, Upper Saddle River: NJ.
63. Tu, X.V. and Dubuisson, B. (1990). 3-D Information Derivation from a pair of binocular images. *Pattern Recognition*, 23, 223-235.
64. Umbaugh, S.E. (1998). *Computer Vision and Image Processing: A Practical approach using CVIPtools*. Prentice-Hall, Eaglewood Cliffs: NJ.
65. Wang, C.Y., Sun, H., Yada, S. and Rosenfeld, A. (1983). Some experiments in relaxation image matching using corner features. *Pattern Recognition*. 16(2), 167-182.

